Long noncoding RNAs are critical regulators of pancreatic islet development and function

Ruth A. Singer

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2019

ABSTRACT

Long noncoding RNAs are critical regulators of pancreatic islet development and function

Ruth A. Singer

Diabetes is a complex group of metabolic disorders with genetic, immunological, and environmental etiologies. Decades of diabetes research have elucidated many genetic drivers of normal islet function and dysfunction. Furthermore, genome wide associated studies (GWAS) have discovered that most diabetes susceptibility loci fall outside of coding regions, which suggests a role for noncoding elements in the development of disease. This highlights our incomplete understanding of the islet regulome and suggests the need for detailed functional analyses of noncoding genes to precisely determine their contribution to diabetes susceptibility and disease progression. Transcriptome analyses have revealed that the eukaryotic genome is pervasively transcribed. Strikingly, only a small proportion of the transcriptome is subsequently translated into protein; the majority is made up non-protein coding RNAs (ncRNAs). The most abundant class of these ncRNAs are called long noncoding RNAs (lncRNAs), defined as transcripts longer than 200 nucleotides that lack protein-coding potential. The establishment of lncRNAs, once dismissed as genomic dark matter, as essential gene regulators in many biological processes has redefined the central role for RNA in cells. While evidence suggests a role for lncRNAs in islets and diabetes, in vivo functional characterization of islet lncRNAs is lacking.

For my thesis project, I sought to understand the lncRNA regulatory mechanisms that promote pancreas development and function. We conducted comparative transcriptome analyses between

embryonic mouse pancreas and adult mouse islets and identified several pancreatic lncRNAs that lie in close proximity to essential pancreatic transcription factors. One of the candidate lncRNAs, *Pax6 Upstream Antisense RNA* (*Paupar*), mapped near *Pax6*, a gene encoding an essential pancreatic regulatory protein. We demonstrate *Paupar* is enriched in glucagon-producing alpha cells where it promotes the alternative splicing of Pax6 to an isoform required for activation of essential alpha cell genes. Consistently, deletion of *Paupar* in mice resulted in dysregulation of Pax6 alpha cell target genes and corresponding alpha cell dysfunction. These findings illustrate a distinct mechanism by which lncRNAs can contribute to cell-specific regulation of broadly expressed transcription factors to coordinate critical functions within a cell.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

Chapter 1

Chapter 2

Appendix

LIST OF ABBREVIATIONS

3P-seq	Poly(A)-position profiling by sequencing
ADP	Adenosine diphosphate
Ago	Argonaute
alphaTCs	AlphaTC cells
amiRNA	Artificial miRNAs
ANS	Autonomic nervous system
Arx	Aristaless-related homeobox
AS	Alternative Splicing
ASO	Antisense oligonucleotides
ATP	Adenosine triphosphate
BAC	Bacterial Artificial Chromosome
Bhlhe22	Basic Helix-Loop-Helix Family Member E22
CAGE	Cap analysis gene expression
Cck	Cholecystokinin
CDC	Center for Disease Control
CHART	Capture Hybridization Analysis of RNA Targets
ChIP	Chromatin Immunoprecipitation
ChIRP	Chromatin Isolation by RNA Purification
CLIP	Cross-linking immunoprecipitation (CLIP)
CO	Capture oligonucleotides
Cpa	Carboxypeptidase A1
CPAT	Coding Potential Assessment Tool
CRAPome	Contaminant Repository for Affinity Purification
DAPI	4′,6-diamidino-2-phenylindole
Deanr1	Definitive endoderm-associated lncRNA 1
DEG	Differentially expressed gene
Dnmt1	DNA Methyltransferase 1
Dyrk1a	Tyrosine-phosphorylation-regulated kinase 1A
E9.5	Embryonic day 9.5
ELISA	Enzyme-linked immunosorbent assay
EP	Embryonic pancreas
eQTL	Expression quantitative trait loci
Fbw7	F-box and WD repeat domain-containing 7
FFA	Free fatty acid
FISH	Fluorescent in situ hybridization
FNF	Flox-Neo-Flox
Foxa1	Forkhead Box A1
Foxa2	Forkhead Box A1
Foxo1	Forkhead Box O1
FPKM	Fragments per kilobase million

Gapdh	Glyceraldehyde 3-phosphate dehydrogenase
Gas5	Growth Arrest Specific 5
Gcg	Glucagon
Gckr	Glucokinase Regulator Protein
GEO	Gene Expression Omnibus
GIP	Gastric inhibitory polypeptide
Glis3	GLIS Family Zinc Finger 3
Glp-1	Glucagon-like- peptide 1
Glut1	Glucose transporter 1
Glut2	Glucose transporter 2
GREAT	Genomic Regions Enrichment of Annotations Tool
GTF	Gene Transfer Format
GWAS	Genome wide association studies
H3K36me3	Histone 3 Lysine 36 trimethylation
H3K4me3	Histone 3 Lysine 4 trimethylation
Hes1	Hairy And Enhancer Of Split 1
hESCs	Human embryonic stem cells
Hnf1b	Hepatocyte nuclear factor 1B
Hnf6a	Hepatocyte nuclear factor 6A
Hotair	HOX transcript antisense RNA
Hottip	HOXA transcript at the distal tip
IDF	International Diabetes Federation
IDT	Integrated DNA Technologies
Ins	Insulin
Insm1	Insulinoma-associated 1
iPSCs	Induced pluripotent stem cells
Isl1	Insulin gene enhancer protein
ITT	Insulin tolerance test
KATP	ATP-sensitive potassium channel
Kcnj11	Potassium Voltage-Gated Channel Subfamily J, Member 11
KD	Knockdown
KO	Knock out
LNA	Locked nucleic acid
LncRNA	Long noncoding RNA
LncRNA-ROR	lncRNA-regulator of reprogramming
MafA	MAF BZIP Transcription Factor A
MafB	MAF BZIP Transcription Factor B
Malat1	Metastasis Associated Lung Adenocarcinoma Transcript 1
Meg3	Maternally Expressed 3 Gene
Miat1	Myocardial infarction associated transcript 1
MIN6	Mouse insulinoma cell line 6
MIP	Mouse insulin promoter
MiRNA	Micro RNA

Mnx1	Motor Neuron And Pancreas Homeobox 1
MPCs	Multipotent progenitor cells
mRNA	Messenger RNA
MS	Mass Spectrometry
mTOR	Mechanistic Target Of Rapamycin Kinase
NCHS	National Center for Health Statistics
ncRNA	Noncoding RNA
Neat1	Nuclear Paraspeckle Assembly Transcript 1
NeuroD1	Neuronal Differentiation 1
Ngn3	Neurogenin3
NHIS	National Health Interview Survey
NIT-1	NOD mice with insulin-promoter/SV40 T-antigen gene
Nkx2-2	NK2 Homeobox 2
NOD	Non-obese diabetic
Paupar	Pax6 Upstream Antisense RNA
Pax4	Paired box 4
Pax6	Paired box 6
Pdx1	Pancreatic And Duodenal Homeobox 1
Pluto	PDX1 Associated LncRNA
Pol II	RNA polymerase II
PP	Pancreatic polypeptide
Prox1	Prospero Homeobox 1
Ptf1a	Pancreatic transcription factor 1 subunit alpha
PYY	Peptide YY
qPCR	Quantitative polymerase chain reaction
qRT-PCR	Quantitative reverse transcription PCR
RAP	RNA Antisense Purification
REST	RE1-silencing transcription factor
Rfx6	Regulatory Factor X6
RIP	RNA immunoprecipitation
RIP:Cre	Rat Insulin Promoter Cre
RISC	RNA-induced silencing complex
RNA-seq	RNA-sequencing
RNAi	RNA interference
RPKM	Reads per kilobase million
rRNA	Ribosomal RNA
Sencr	Smooth muscle and endothelial cell enriched migration-associated lncRNA
Sglt2	Sodium/glucose co-transporter 2
shRNA	Short hairpin RNA
siRNA	Small interfering RNA
Slc30a8	Solute carrier family 30, member 8
smFISH	Single Molecule FISH
SNP	Single nucleotide polymorphism

Sox6	SRY-Box 6
Sox9	SRY-Box 9
STRING	Search Tool for the Retrieval of Interacting Genes/Proteins
T1DM	Type 1 diabetes mellitus
T2DM	Type 2 diabetes mellitus
TBP	TATA-binding protein
Terc	Telomerase RNA Component
Tgif2	TGFB Induced Factor Homeobox 2
tRNA	Transfer RNA
TSS	Transcriptional start site
Tug1	Taurine Up-Regulated 1
TZD	Thiazolidinedione
VDCC	Voltage-dependent Ca+-channels
Xist	X Inactive Specific Transcript
βlinc1	β-cell long intergenic noncoding RNA 1

ACKNOWLEDGEMENTS

I would like to begin by thanking my mentor, Lori Sussel, for her endless support and guidance. I have had the privilege of knowing Lori for almost 10 years. I first started in her lab as a technician, fresh out of college with almost no lab experience and a tenuous desire to go to medical school. Lori always treated me as more than a technician, giving me multiple opportunities to publish and spearhead my own projects. Lori's love of science is infectious, and with her guidance, I quickly (3 years down the road) decided that biomedical research was my true passion. I entered the graduate school process with no inkling that I wanted to stay at Columbia, or that I would return to Lori's lab for my thesis work. For those who know Lori well, it probably comes as no surprise that I quickly realized how incomparable Lori was as a mentor and that I wanted to do my thesis work in her lab. While her move to Colorado in my third year was quite a curveball, her support for me back in New York never wavered. Whenever the project seemed hopeless, or an experiment seemed like it would never work, I always felt better after a chat (usually over FaceTime) with Lori. When I messed up an experiment, she was kind, encouraging, and quick to remind me "if it was easy, then everyone would do it." I feel very fortunate to have had the opportunity to work in Lori's lab and know I am a better scientist and person because of her.

I also would like to extend a great deal of gratitude to the members of the Sussel lab. The members of Lori's Columbia lab made coming into lab every day so fun. They taught me everything I know about developmental biology, molecular biology, and how to think through my experiments. I would especially like to thank Dina Balderes, for her constant support. As my

bay mate, she was the first person I went to when experiments failed. She always helped me talk through what didn't work and would help me figure out how to get it right the next time. Her love of science was infectious, and I am so grateful for how she kept our lab running smoothly. To the Colorado Susselites, thank you for welcoming me so earnestly into your crew. When I visited the lab, you all went out of your way to make me feel welcome, from taking the day off to go hiking to letting me tag along on family trips to the local pool. I would like to give David a special shout out for bringing me to all the lab meetings on your computer via FaceTime.

I would also like to thank the CMBS class of 2013. You all are such incredible people and scientists and I am so lucky to have gotten to go through this whole process with you guys by my side. I would especially like to thank Chelsea, Margot, and Clau for countless wine nights that kept me going through all these years. Thank you to the incredible labs of the Russ Berrie Diabetes Center, especially the Egli, Qiang, and Leibel labs that treated me as one of their own and provided me with many reagents over the years. I am especially grateful to Hector in the mouse facilities for taking such good care of my mice and giving me a heads up via text when I had several overcrowded cages.

I would like to thank my incredible family. My parents, Miriam and Steve, are my biggest cheerleaders and I literally owe everything I have to them. Thank you to my sisters and my amazing brothers-in-law for keeping me on my toes and being my favorite people. Finally, I would like to thank my amazing husband, partner, and scientific advisor, Ari. Thank you for unwavering support and for giving meaning to everything in my life. I could not have done this without you.

DEDICATION

I dedicate this work to my family, my incredible husband, the soon-to-be Dr. Ari Zolin Ph.D., and my darling 7-month-old niece Noa, who represents all the fearless women who will change our world for the better.

CHAPTER 1: INTRODUCTION

I. Pancreatic islet development

Pancreas anatomy

The pancreas is a dual functioning organ that plays a critical role in the regulation of macronutrient digestion and blood glucose levels. The majority of the pancreas is made up of exocrine, or acinar, cells that secrete digestive enzymes, such as amylase, lipase, trypsin, and chymotrypsin, into the duodenum via the pancreatic ducts (Figure 1-1). In contrast, pancreatic hormones are released in an endocrine manner via direct secretion into the blood stream (Figure 1-1). The endocrine cells, which account for only 1–2% of the entire organ, are clustered together in island-like structures, aptly named islets of Langerhans, within the exocrine pancreatic tissue (Figure 1-1). There are four different cell types releasing various hormones from the pancreatic endocrine compartment: glucagon-producing alpha cells (15-20% of mouse islets; ~40% of human islets), insulin-producing beta cells (60-80% of mouse islets, ~50% of humans islets), somatostatin-producing delta cells (<10% of both mouse and human islets), and pancreatic polypeptide-producing PP-cells (<2% of both mouse and human islets) (Steiner et al., 2011; Brissova et al., 2005; Cabrera et al., 2006) (Figure 1-2). Two additional endocrine cell types, gastrin-producing G cells (Suissa et al., 2013) and ghrelin producing epsilon cells (Prado et al., 2004; Sussel et al., 1998), are found transiently in the embryonic pancreas. Interestingly, in most mammalian species, the different islet endocrine cells are organized in a non-random pattern, with a core of beta cells surrounded by a discontinuous mantle of delta, alpha, and pancreatic polypeptide cells (Erlandsen et al. 1976; Orci and Unger 1975; Shih et al., 2013) (Figure 1-2). In contrast, several aspects of human islet morphology have remained

controversial, including islet composition and islet architecture. Recently, studies using confocal microscopy concluded that human islets lack the typical core-mantle structure and non-beta cells were dispersed throughout the islet (Brissova et al., 2005; Cabrera et al. 2006). However, a comprehensive study by Bonner-Weir and colleagues using non-diabetic adult human pancreas and enhanced microscopy technologies showed that there was far more variability in islet composition in human pancreas than rodent pancreas (Bonner-Weir et al., 2015). They concluded that while a subset of human islets had interspersed alpha and beta cells within the islet core, the fundamental architectural arrangement is similar between mice and men (Bonner-Weir et al., 2015).

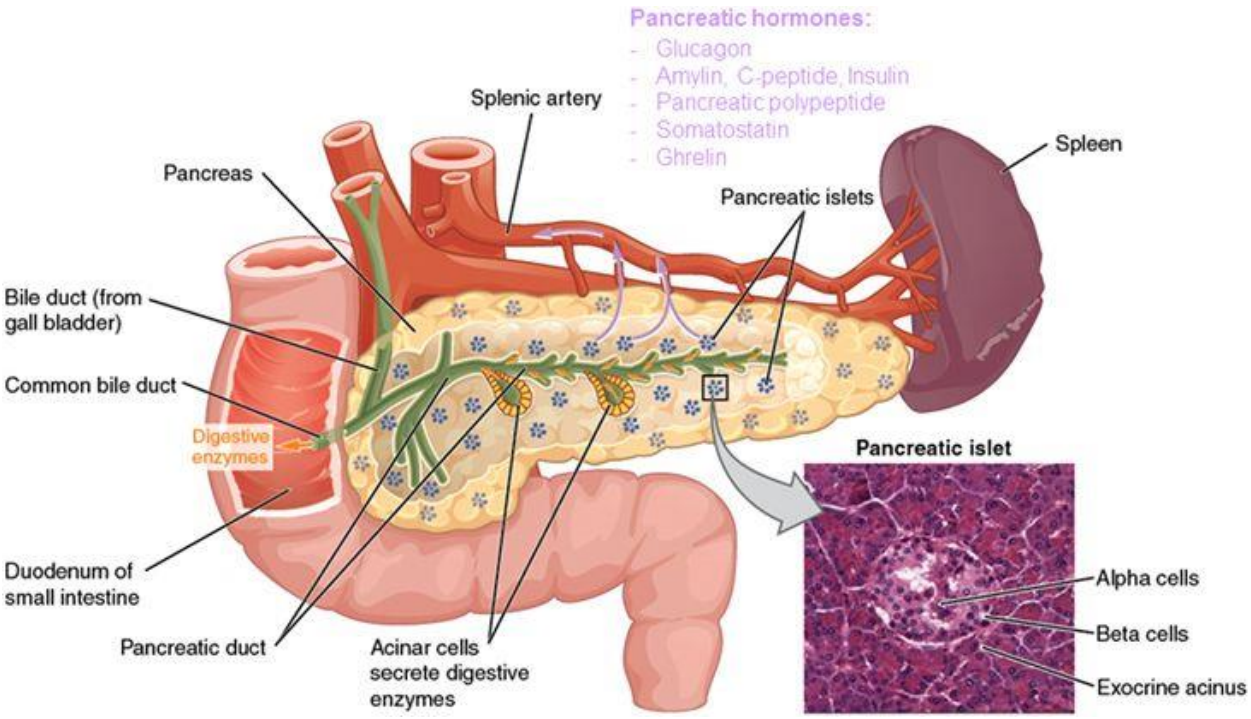


Figure 1-1. Anatomical organization of the pancreas. The exocrine function of the pancreas is mediated by acinar cells that secrete digestive enzymes into the upper small intestine via the pancreatic duct. Its endocrine function involves the secretion of various hormones from different cell types within the pancreatic islets of Langerhans. The micrograph shows the pancreatic islets. Adapted from Roder et al., 2016.

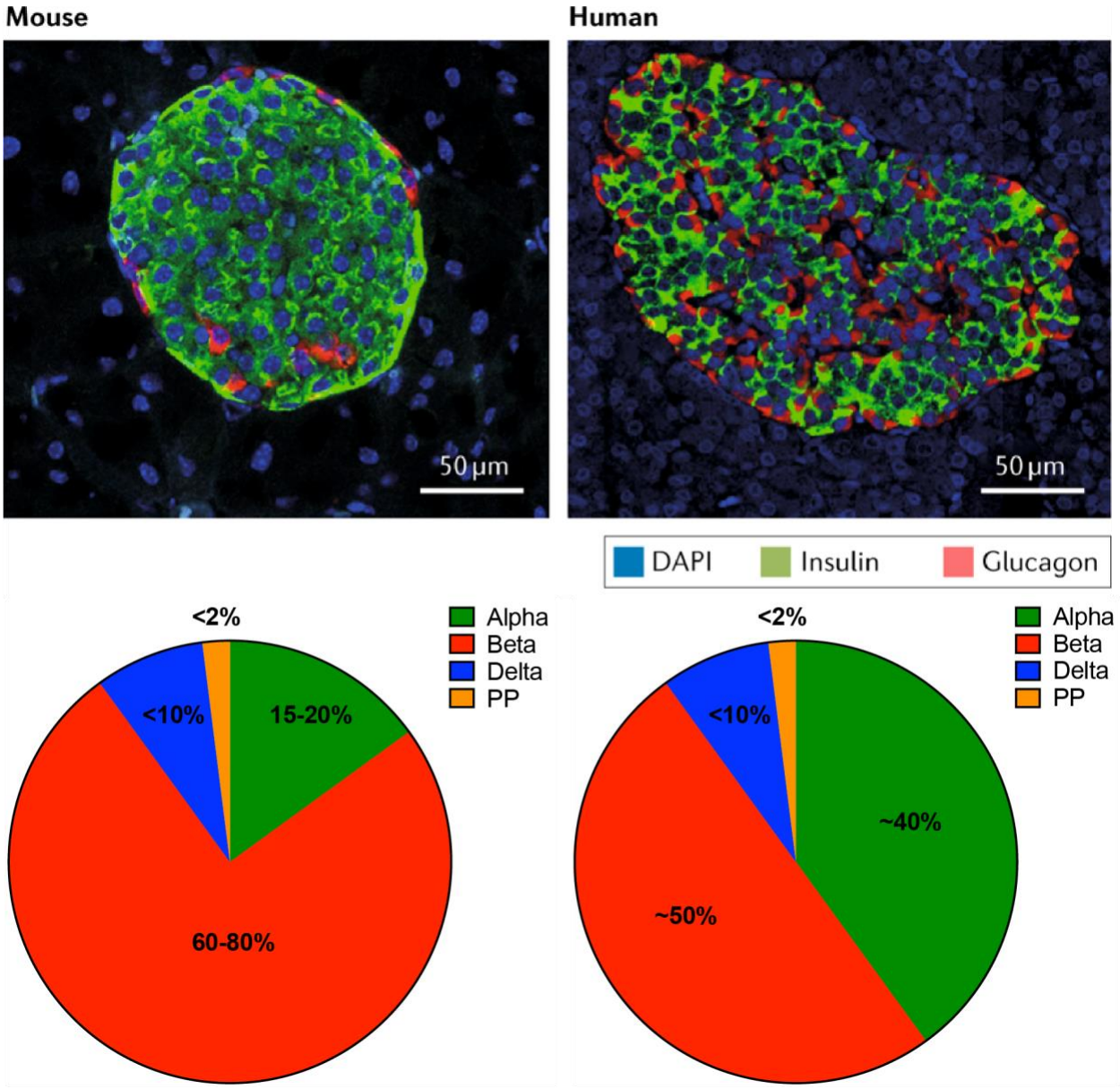


Figure 1-2. The localization and number of alpha cells differ between mouse and human pancreatic islets. Mouse islets have a beta cell-rich core (insulin stain in green), which is surrounded by a low number of alpha cells (glucagon stain in red). Human islets are composed of substructures, each with an arrangement similar to that of mouse islets. Cell nuclei are stained in blue. The pie charts show the proportions of the major endocrine cell types in mouse and human islets. DAPI, 4′,6-diamidino-2-phenylindole; PP, pancreatic polypeptide. Adapted from Gromada et al 2018.

Specification of pancreatic progenitor cells

The development of the pancreas is a complex process in which two morphologically distinct tissue types, exocrine and endocrine cells, must derive from one simple epithelium (Gu et al., 2002). In order to form the mature architecture of the pancreas, there is a carefully orchestrated series of events mediated by extrinsic signals from adjacent mesodermal derivatives, as well as by intrinsic programs controlled by factors expressed within the endodermal cells themselves (reviewed in Gittes, 2009; Jørgensen et al., 2007; Larsen and Grapin-Botton, 2017; Shih et al., 2013; Pan and Wright, 2011). Mouse pancreas development begins at embryonic day 9.0 (E9.0) when the dorsal foregut endoderm thickens and evaginates into the surrounding mesenchyme (Munger, 1958; Wessells and Cohen, 1967; Kallman and Grobstein, 1964; Pictet et al., 1972) (Figure 1-3). Approximately 12 hours later, the ventral pancreas and common bile duct emerge from the ventral foregut endoderm. At E11.5, the gut tube begins to undergo rotation and elongation that allows the ventral and dorsal buds to come into contact with one another. Subsequent fusing of the dorsal and ventral buds into a single anlage occurs at ~E12 (Figure 1-3). This process is not well understood, but when it is disrupted as in hedgehog mutants, malformations such as annular pancreas (a ring of pancreatic tissue that completely encircles the duodenum) can develop (Ramalho-Santos et al., 2000; Hebrok et al., 2000). Interestingly, at this early stage, isolated pancreatic endoderm can grow and differentiate in vitro to form pancreatic structures in organ culture experiments (Wessells and Cohen, 1967). However, growth and differentiation at this stage is dependent on signaling factors from the pancreatic mesenchyme, which was demonstrated by elegant co-culture experiments where pancreas development could proceed although the epithelium was separated from the mesenchyme by a porous membrane (Golosow and Grobstein, 1962).

Differentiation of pancreatic progenitor cells

During this early phase of pancreatic development, which has been referred to as the primary transition (Pictet et al., 1972), a complex network of transcriptional regulators mediates expansion of pancreatic progenitor cells and maintains pancreatic identity over other lineages (Figure 1-3). The pancreatic and duodenal homeobox 1 (Pdx1) transcription factor is the earliest detectable marker of the pancreatic anlage and is one of the most central nodes in the pancreatic transcriptional network (Ahlgren et al., 1996; Guz et al., 1995) (Figure 1-4). Genetic lineage tracing experiments showed that Pdx1+ cells represent progenitors of all the mature pancreatic cell types, including duct, islet, and acinar cells (Gu et al., 2002). Mice deleted for Pdx1 exhibit complete pancreatic agenesis at birth despite initial formation of epithelial buds (Offield et al., 1996; Marty-Santos and Cleaver, 2016; Ahlgren et al., 1996). Furthermore, maintenance of the pancreatic lineage requires higher levels of Pdx1 (Offield et al., 1996; Fujitani et al., 2006), which appears to be partly achieved by binding of pancreatic transcription factor 1 subunit alpha (Ptf1a) to enhancer elements upstream of Pdx1 (Wiebe et al., 2007). Ptf1a expression is initiated after Pdx1 in the posterior foregut endoderm (Burlison et al., 2008) and has a critical function in pancreas development (Obata et al., 2001; Kawaguchi et al., 2002; Krapp et al., 1998; Sellick et al., 2004; Weedon et al., 2014). In addition to Pdx1 and Ptf1a, several other transcription factors have been shown to be important for the primary transition, including Sox9, Gata4/6, Foxa1/2, Hnf6a (Onecut1), Hnf1b (Tcf2, Onecut2), Prox1, and Mnx1 (Ahlgren et al. 1996, Guz et al. 1995, Kawaguchi et al. 2002, Krapp et al. 1998, Seymour et al. 2007) (Figure 1-4). Importantly, mice lacking any one of these factors display varying degrees of pancreas hypoplasia or agenesis (reviewed in Gittes, 2009, Pan and Wright, 2011, Seymour and Sander, 2011).

Specification of pancreatic endocrine progenitor cells

Shortly after the primary transition, the nascent pancreatic buds are almost entirely composed of multipotent progenitor cells (MPCs). During the secondary transition (E13.5-E15.5), non-committed multipotent progenitor cells (MPCs) differentiate into an organ with different cellular compartments; the pancreatic epithelium undergoes dynamic structural changes, resulting in multiple protrusions that bud from the edges (Figure 1-3). These structures consist of a tip domain, marked by expression of Ptf1a and Cpa, and a trunk domain, identified by Nkx6.1/6.2, Sox9, Hnf6b, Hnf6a, Prox1, and Hes1 (Schaffer et al. 2010; Klinck et al., 2011; Kopinke et al. 2011, Jacquemin et al. 2003, Vanhorenbeeck et al. 2007, Wang et al. 2005, Zhou et al. 2007) (Figure 1-4). Lineage-tracing experiments revealed that the trunk predominantly gives rise to the endocrine and ductal cell lineages, whereas the tips are quickly restricted to an acinar fate (Kopinke et al. 2011, Kopp et al. 2011, Pan et al. 2013, Solar et al. 2009, Zhou et al. 2007) (Figure 1-3). The transcription factors Nkx6.1 and Ptf1a act as master regulators during this process. These factors are initially co-expressed in MPCs, but their expression domains entirely segregate during tip and trunk compartmentalization: Nkx6.1/6.2 promote trunk identity and segregate to trunk cells, and Ptf1a has an equivalent role in the tip compartment (Schaffer et al. 2010). Mechanistically, this process is initiated by transcriptional cross-repression between Nkx6.1/6.2 and Ptf1a (Schaffer et al. 2010) (Figure 1-4).

Downstream of Nkx6.1-mediated specification of trunk cells, several factors play a role in the ductal versus endocrine fate decision. The central intrinsic regulator of endocrine specification is the bHLH transcription factor Ngn3, the expression of which is necessary and sufficient for endocrine development in the pancreatic endoderm (Gu et al., 2002; Apelqvist et al., 1999;

Gradwohl et al., 2000; Grapin-Botton et al., 2001; Schwitzgebel et al., 2000) (Figure 1-4). Mice lacking Ngn3 function fail to develop endocrine cells (Gradwohl et al., 2000), whereas ectopic Ngn3 expression under the Pdx1 promoter induces premature differentiation of the entire pancreatic bud progenitor pool into endocrine cells (Apelqvist et al., 1999; Schwitzgebel et al., 2000). Surprisingly, Ngn3 is expressed in only a subset of cells within the trunk domain, which raises the question of how its expression is repressed in the majority of embryonic ductal cells. Biochemical and genetic evidence suggests that the transcription factor Hes1 plays an important role in repressing Ngn3 transcription and preventing widespread Ngn3 activation (Ahnfelt-Ronne et al. 2012, Apelqvist et al. 1999, Jensen et al. 2000, Lee et al. 2001). Recently, Hes1 was found to accelerate NGN3 protein degradation (Qu et al. 2013), which suggests that Ngn3 is regulated posttranslationally. Supporting this idea, a separate study reported that Ngn3 mRNA is detected in a much broader domain than NGN3 protein (Villasenor et al., 2008).

Differentiation pancreatic endocrine progenitor cells

Once Ngn3 is expressed, progenitors exit the cell cycle, delaminate, and migrate away from the progenitor cords to form endocrine clusters (Gouzi et al. 2011, Miyatsuka et al. 2011). During this process, endocrine precursors go through progressive differentiation towards mature endocrine cells by initiating temporal waves of transcription factors ensuring unidirectional differentiation. It is important to note that some transcription factors promote cellular lineage decisions by activating downstream genetic targets, while others act as repressors of alternative cellular programs. Ngn3 activates the expression of a number of endocrine transcription factors including: NeuroD1, Pax4, Arx, Insm1, Rfx6, Nkx2.2, and Myt1 (Pan and Wright, 2011; Huang et al., 2000; Smith et al., 2003; Smith et al., 2004; Watada et al., 2003; Mellitzer et al., 2006;

Wang et al., 2007; Rukstalis and Habener, 2007; Soyer et al., 2010; Larsen and Grapin-Botton 2017) (Figure 1-4). Isl1 is also thought to be downstream of Ngn3 (Du et al., 2009), but whether Ngn3 directly activates Isl1 expression remains to be determined. Eventually, the combinatorial expression of multiple downstream factors will define the identity of specific islet cell types (reviewed in Spence and Wells, 2007). Over the next several days (E14-E18) the maturing endocrine cells begin to accumulate along the ducts and blood vessels in a cord-like linear pattern. In the neonatal pancreas, these linear endocrine collections begin to coalesce into aggregates that represent the first islets of Langerhans (Figure 1-3). Although conventional lineage diagrams depict single *Ngn3+* cells giving rise to multiple islet cell types, it is equally plausible that four (or more) independent classes of *Ngn3+* cells exist, one for each mature cell type in the islet (Murtaugh, 2007). Distinguishing these models, however, will require lineage-tracing techniques of considerable higher resolution than are currently available (Murtaugh, 2007).

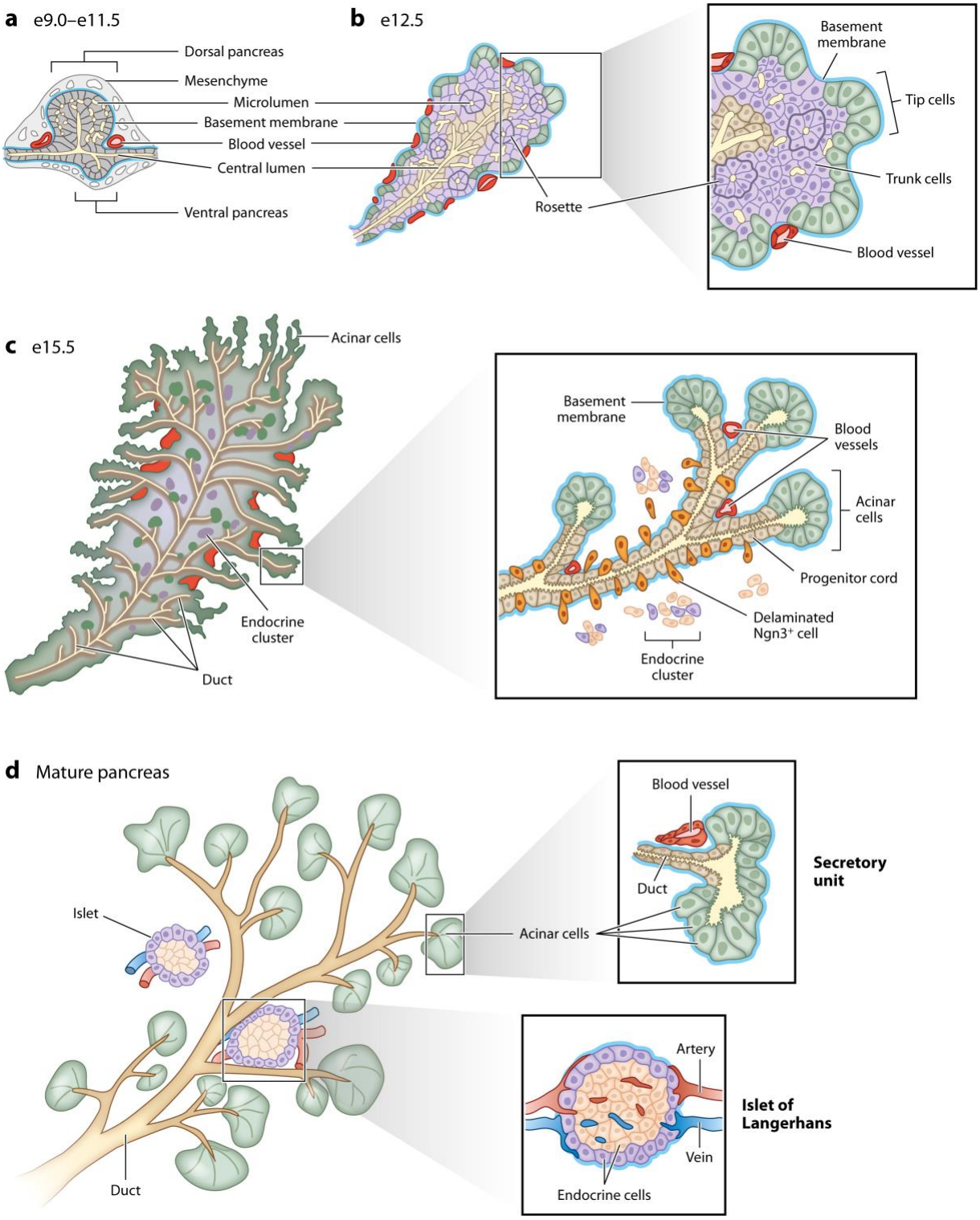


Figure 1-3. Illustrated overview of pancreatic organogenesis. (A) The dorsal and ventral pancreatic epithelium evaginates into the surrounding mesenchyme between e9.0-e11.5 in mice.

At this stage, the pancreatic epithelium is comprised of a multilayered core of unpolarized cells engulfed by a basement membrane. Scattered microlumens (light yellow) arise between epithelial cells. Blood vessels surround but have not yet penetrated the epithelial buds. (B) At e12.5, the outer tip cell layer (green) of the pancreatic epithelium forms recognizable branch protrusions. In the trunk portion of the pancreatic epithelium, microlumens fuse to form a primitive plexus, and groups of newly polarized cells (purple) organize into rosettes around a lumen. At the same time, blood vessels begin to intercalate into the epithelium and contact trunk cells. (C) At e15.5, the luminal plexus progressively remodels into a single-layered epithelium consisting of highly branched primitive ducts (also known as progenitor cords) and newly differentiated acinar cells. Ngn3-expressing endocrine precursors (orange) delaminate and migrate away from the progenitor cords to form endocrine clusters. Blood vessels are intercalated between nascent branches of the pancreatic ductal tree. (D) In the mature pancreas, acinar cells cap the endings of small terminal ducts and form functional exocrine secretory units. Endocrine cells are clustered in so-called islets of Langerhans, which are penetrated by a dense network of blood vessels. Adapted from Shih et al., 2013

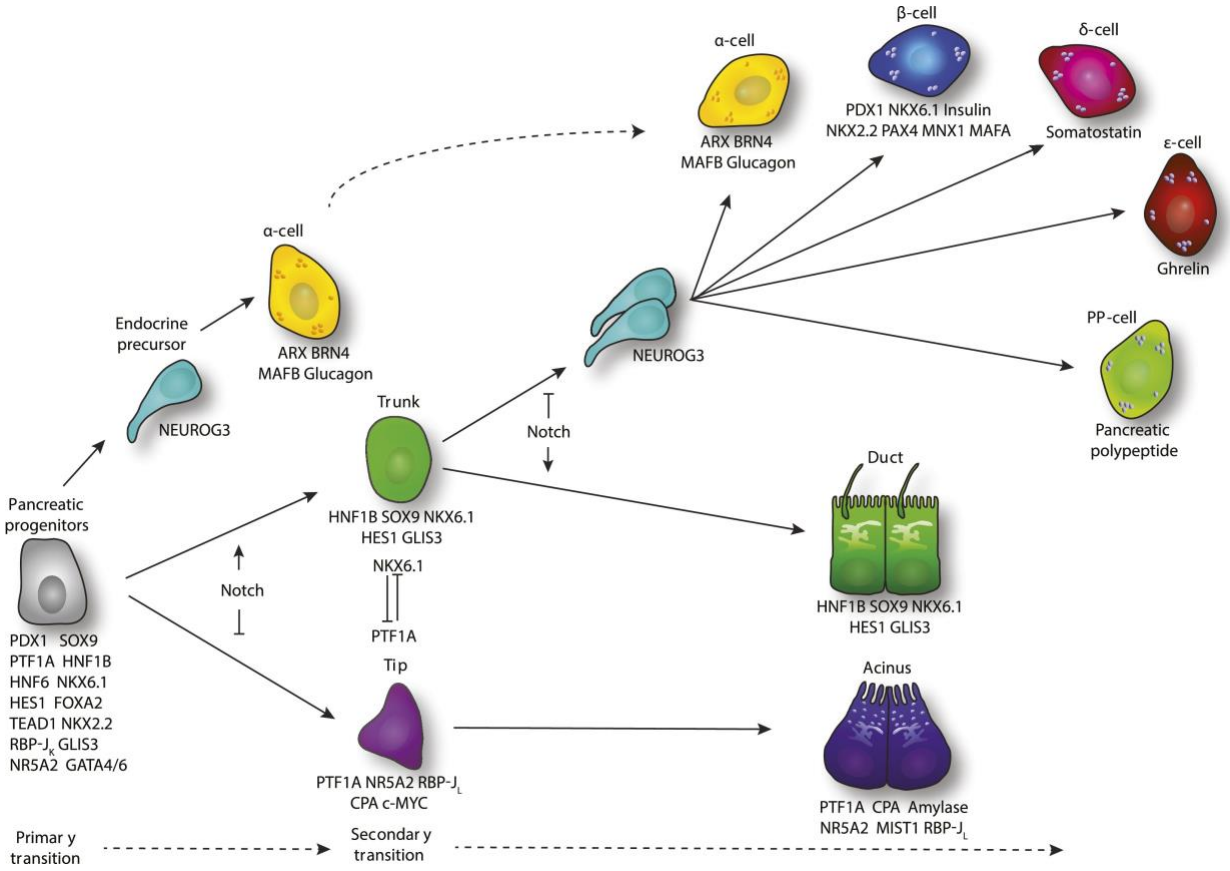


Figure 1-4. Lineage hierarchy during pancreas organogenesis. Following specification, the multipotent pancreatic progenitors co-express a range of pancreas-associated transcription factors facilitating establishment of the gene regulatory network mediating multipotency at the early stages of pancreas development. Priming towards the endocrine lineage occurs in scattered multipotent progenitors via transient expression of Ngn3, subsequently leading to the emergence of primary transition-derived alpha cells. Notch signaling and reconfiguration of the transcriptional connectivity between Ptf1a and Nkx6.1 next initiate tip-trunk segregation, leading to the segregation of the acinar and ducto-endocrine lineages. In the ductal-endocrine bipotent population, scattered progenitors are primed towards the endocrine lineage via NGN3 expression facilitated by low levels of Notch signaling. The population of secondary transition-derived endocrine precursors primarily gives rise to beta cells but also to the other four endocrine

subtypes. Remaining ductal progenitors eventually mature into ductal cells displaying primary cilia and hydrogen bicarbonate production. Adapted from Larsen and Grapin-Botton, 2017.

Alpha and Beta cell specification and maturation

Endocrine lineage specification towards mono-hormonal endocrine cells is controlled by the expression of lineage-specific transcription factors (Figure 1-4). For example, the transcription factors Pax4, Nkx2-2, Pdx1, and Nkx6.1 are all critical beta cell determinants (Sussel et al., 1998; Collombat et al. 2003, Gannon et al. 2008, Henseleit et al. 2005, Holland et al. 2002, Sosa-Pineda et al. 1997) (Figure 1-5). Similarly, several essential alpha cell transcription factors have been identified, including Pax6, Arx, and Foxa2, and mice lacking any of these factors do not produce functional alpha cells (Gromada et al., 2007; Sander et al., 1997; Masson et al., 2014; Jin, 2008) (Figure 1-5). Following the initial instigation of endocrine subtype lineages, subsequent functional maturation is orchestrated by another set of transcription factors. The transcription factors MafA and MafB play pivotal roles in the functional maturation of beta and alpha cells, respectively. MafB is expressed in nascent embryonic alpha and beta cells and becomes restricted to mature alpha cells where it directly regulates the expression of glucagon by binding to upstream promoter elements (Artner et al., 2006). While immature beta cells express MafB, fully mature beta cells are characterized by the expression of MafA and absence of MafB (Artner et al., 2007; Artner et al., 2010; Nishimura et al., 2006), although this switch is not conserved in humans (Riedel et al., 2012). Expression of MafA is controlled by direct transcriptional activation of beta cell factors such as Pdx1, Nkx2.2, Foxa2 and Pax6 (Raum et al., 2006; Zhang et al., 2005).

Cell specific function of pancreatic transcription factors

Studies in mice have shown that several pancreatic transcription factors have dual regulatory roles: a single transcription factor can be required during development for cell lineage decisions

and for the maintenance of mature islet cell identity and function. There are several examples of these dual regulators in the literature. Nkx6-1 has been shown to be necessary and sufficient to promote differentiation of endocrine progenitor cells into mature beta cells (Schaffer et al., 2013). Once the beta cells have matured, Nkx6-1 expression must be maintained or the cells will acquire molecular characteristics of delta cells, initiating the rapid onset of diabetes (Taylor et al., 2013) (Figure 1-5). This archetype is not unique to Nkx6-1; multiple transcription factors have been shown to be required during both pancreas development and in mature beta cells, including Nkx2-2 (Sussel et al., 1998; Gutierrez et al., 2016; Churchill et al., 2016), NeuroD1 (Anderson et al., 2009; Naya et al., 1997), Glis3 (Kang et al., 2016), Pdx1 (Offield et al., 1996; Gao et al., 2014) (Figure 1-5). This archetype of transcription factors is also not unique to beta cells. The transcription factor, Arx, has been shown to be required for alpha cell development (Collombat et al., 2003) and the maintenance of mature alpha cell identity and function (Courtney et al., 2013) (Figure 1-5).

To make matters even more complicated, some transcription factors have triple duty: they are required during development and in several different mature islet cell types. One of the more characterized pancreatic transcription factors, Pax6, exhibits these regulatory properties. Pax6 is required downstream of Ngn3 for proper development of all endocrine lineages, hormone production and islet organization (Sander et al., 1997; St-Onge, 1997). Conditional deletion of Pax6 in mature alpha or beta cells showed that Pax6 is also required to maintain the identity and function of both endocrine cell types (Ahmad et al., 2015; Gosmain et al., 2012; Mitchell et al., 2017; Swisa et al., 2017). Some cell specific regulatory function of Pax6 has been explained; in alpha cells, Pax6 heterodimerizes with MafB and stimulates *Gcg* expression by binding to the G1

element (Gauthier et al., 2002; Gosmain et al., 2007) (Figure 1-6); in beta cells, Pdx1, Pax4, and Nkx6.1 bind to G1 and inhibit *Gcg* expression by blocking the binding of the Pax6/Maf heterodimer to the G1 element (Gauthier et al., 2007, Gosmain et al., 2007; Ritz-Laser et al., 2002) (Figure 1-6). Although these studies shed light on how the *Glucagon* gene is activated in alpha cells and repressed in beta cells, many other cell-specific, or gene-specific, transcription factor regulatory mechanisms remain poorly understood. For example, recent work has shown that within mouse beta cells, Pax6 directly activates critical beta cell genes and represses genes that specify the alternative islet endocrine cell lineages (Swisa et al., 2017). Similar findings were reported for the transcription factor, Pdx1, which acts as an activator of mature beta cell genes (Gao et al., 2014). Interestingly, conditional deletion of Pdx1 in mature beta cells induced upregulation the alpha cell transcription factor, MafB (Gao et al., 2014). This study further demonstrated that Pdx1 normally binds and represses MafB beta cells, showing that Pdx1 has both activating and repressive functions, depending on the cellular context (Gao et al., 2014). I will describe the regulatory role of Pax6 in islets in greater detail in Chapter 2, but collectively, these findings highlight that the molecular mechanism mediating cell specific regulatory functions of transcription factors require further investigation.

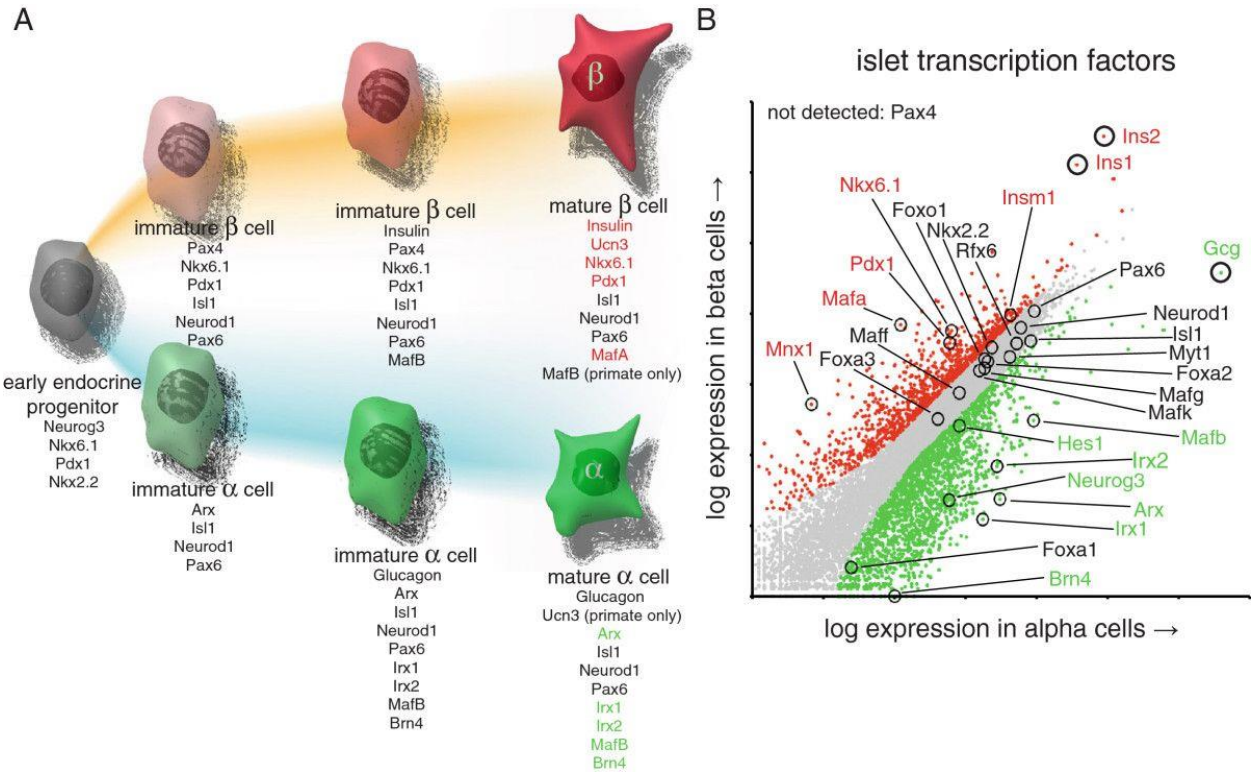


Figure 1-5. Comparison of mouse beta and alpha cell transcriptomes based on functional annotation. Establishment and maintenance of alpha and beta cell identity is regulated by a complex interplay of transcription factors (A), whose expression pattern is accurately reflected by our mouse alpha and beta cell transcriptomes. A dot plot represents actual transcription factor expression in beta and alpha cells where each dot represents an individual gene (B). Genes that are significantly enriched in beta or alpha cells are highlighted in red and green, respectively. Adapted from Benner et al., 2014.

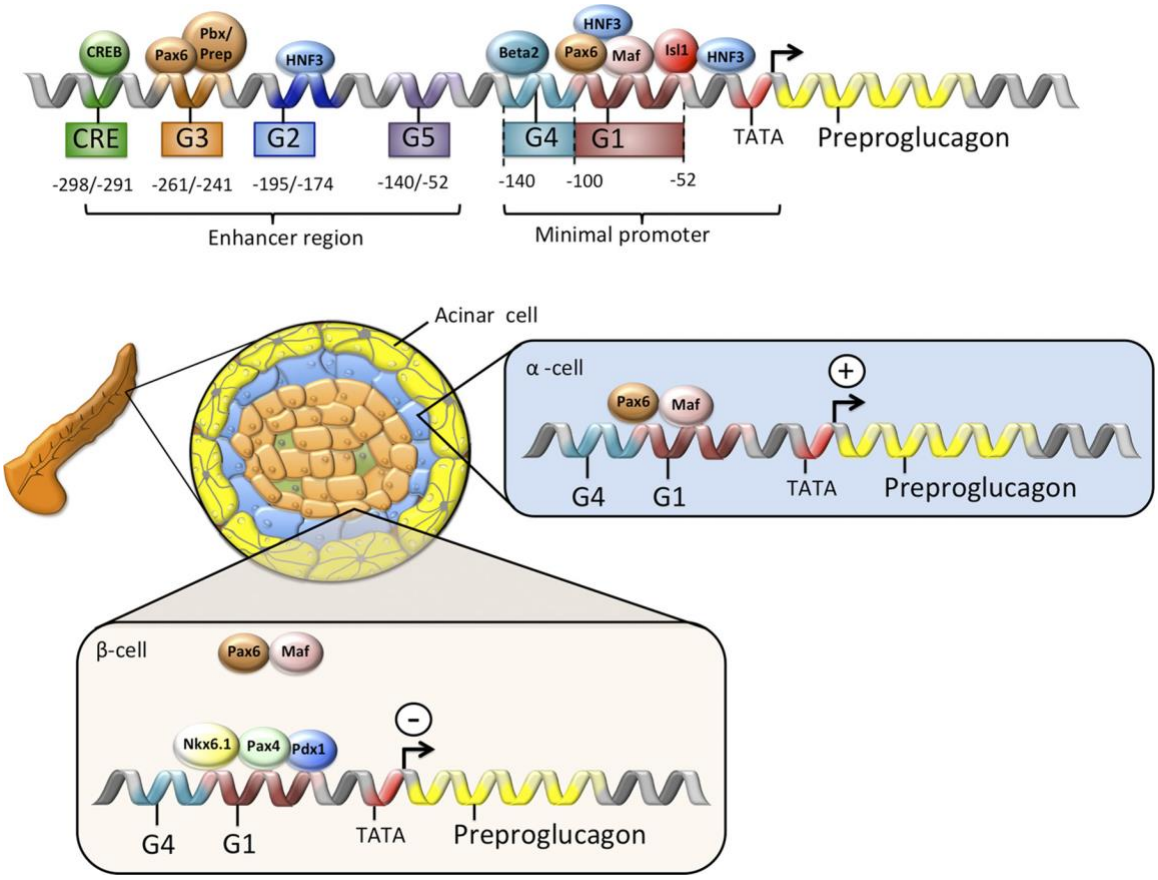


Figure 1-6. Schematic on the transcriptional regulation of preproglucagon in the pancreatic alpha and beta cells. The expression of preproglucagon is regulated through interaction of homeodomain proteins that bind to the preproglucagon promoter region, which comprises a minimal promoter region and an enhancer region. Adapted from Muller et al., 2017.

II. Pancreatic islet function

Bihormonal regulation of glucose homeostasis

Each islet hormone has a distinct function: insulin decreases blood glucose levels (Goke 2008); glucagon increases blood glucose levels; somatostatin inhibits glucagon and insulin release (Hauge-Evans et al., 2009); PP regulates the exocrine and endocrine secretion activity of the pancreas (Katsuura et al., 2002). Together, these hormones maintain blood glucose levels within a very narrow physiological range (4-6 mM), referred to as glucose homeostasis (Figure 1-7). Briefly, during sleep or between meals, when blood glucose levels are low, glucagon is released from alpha cells to promote hepatic glycogenolysis and gluconeogenesis thereby increasing endogenous blood glucose (Freychet et al., 1988). In contrast, postprandial hyperglycemia stimulates insulin secretion from beta cells (Komatsu et al., 2013). Insulin acts on its receptors in muscle and adipose tissue to promote the uptake of glucose and thereby lowers blood glucose levels by removing the exogenous glucose from the blood stream (Khan and Pessin, 2002, Kohn et al., 1996, Zisman et al., 2000). In these next two sections, I will explore the pivotal findings that have led to our understanding of the role of insulin and glucagon in regulating glucose homeostasis (Figure 1-8).

Beta cells and insulin

The discovery of insulin was one of the key achievements of the twentieth century and arguably the first instance in which science provided a new lifesaving medicine for the morbidly ill. Prior to the discovery of insulin, a diagnosis of diabetes, particularly in children, was equivalent to a death sentence (Bliss, 1982). The only existing treatment was a starvation diet advocated by Frederick Madison Allen, a physician and diabetes expert working at the Rockefeller Institute for

Medical Research (now Rockefeller University) (Allen, 1913). In 1889, Oskar Minkowski and Josef von Mering observed that removal of a dog's pancreas, but not ligation of the pancreatic duct, resulted in polyuria, polydipsia, and diabetes, suggesting that the pancreas is of crucial importance for maintaining euglycemia (von Mering and Minkowski, 1890). Another study noted that the islets of Langerhans in the pancreas were often destroyed in humans with diabetes, but how this pancreatic component controlled sugar metabolism was unknown (Opie, 1900). One possibility was that the islets of Langerhans produced a hormone, referred to then as an "internal secretion," that regulated glucose concentrations in the blood. In the years that followed, work from several scientists, including Israel Kleiner (Kleiner, 1919), Georg Ludwig Zuelzer (Zuelzer, 1908), and Nicolae Paulescu (Paulescu, 1920), successfully used pancreatic extracts to lower blood glucose in diabetic dogs. These studies collectively paved the way for the work of Frederick Grant Banting (Friedman, 2010). In 1921, Banting approached John James Rickard Macleod, a physiology professor at the University of Toronto, and proposed a study to eliminate the exocrine part of the pancreas and enrich for pancreatic islets (Ceranowicz et al., 2015). Macleod accepted Banting's proposal and, before leaving on summer break, gave him lab space, ten dogs, and the help of two medical students, Charles Herbert Best and E. Clark Noble (Wright, 2002). Banting felt he only needed one extra pair of hands, and Best and Noble decided to toss a coin. Best won the coin toss and Banting and Best began a series of experiments that led to the discovery of insulin (Banting and Best, 1922). They initially sourced insulin from dog pancreases, then from pancreases harvested from cattle embryos, and finally from ox pancreases. After Macleod returned from vacation he assumed oversight over the investigations. The alcohol extract that Banting obtained from pancreases lowered glucose levels in dogs that had induced diabetes, but because the extracted insulin was relatively impure, it was unsuitable for clinical

studies, which is why an experienced biochemist, James Betram Collip, joined the team in December 1921. Collip developed a successful method that enabled the extraction of large amounts of highly purified insulin from ox pancreases, and this made it possible for a clinical study to commence in January 1922. The study was a great success and confirmed insulin's effectiveness in treating diabetes. In 1922, Banting, Best, and Collip were granted patent rights for producing insulin, but they sold them to the University of Toronto for a dollar (Friedman, 2010). In 1923, Frederick Grant Banting and John James Rickard Macleod were awarded the Nobel Prize in recognition of their discovery. The first study to suggest diversity among pancreatic islet cells identified two populations of cells, the A cells and B cells (Lane 1907). In B cells, which are now known as beta cells, researchers found basophilic granules that were not present in A cells (Stanger and Samols, 1991). In 1957, a study used immunohistochemistry to show that insulin was the hormone produced by beta cells (Lacy and Davies, 1957). Later, researchers used electron microscopy on human beta cells to determine the ultrastructure of beta cells (Like, 1967).

The primary functional feature of the beta cell is its ability to secrete insulin in a tightly regulated manner in response to metabolic demands. There are different molecules that can promote insulin secretion, such as the glucagon-like- peptide 1 (GLP-1), gastric inhibitory polypeptide (GIP), cholecystokinin (CCK), and peptide YY (PYY). However, the main stimulus for insulin release is elevated blood glucose levels following a meal (Poitout et al., 2006). In beta cells, glucose is taken up by the facilitative glucose transporter, Glut2 (Thorens et al., 1988). Once inside the cell, glucose undergoes glycolysis resulting in increased cellular ratios of free adenosine triphosphate (ATP) to adenosine diphosphate (ADP) (Kennedy et al., 1999). This

altered ratio then leads to the closure of ATP-sensitive K+-channels (KATP) (Ashcroft et al., 1984; Cook and Hales, 1984). Under non-stimulated conditions, these channels are open to ensure the maintenance of the resting potential by transporting positively charged K+-ions down their concentration gradient out of the cell. Upon closure, the subsequent decrease in the magnitude of the outwardly directed K+-current elicits the depolarization of the membrane, followed by the opening of voltage-dependent Ca+-channels (VDCCs) (Roder 2016). The increase in intracellular calcium concentrations eventually triggers the fusion of insulin-containing granules with the membrane and the subsequent release of their content (Jonkers and Henquin 2001; Yang and Berggren, 2006). The insulin secretory process is biphasic with the first phase peaking around 5 minutes after the glucose stimulus with the majority of insulin being released during this first phase. In the second, somewhat slower, phase, the remaining insulin is secreted (Ceraso and Luft, 1967; Porte and Pupo, 1969; Curry et al, 1968).

Alpha cells and glucagon

The monumental discovery of insulin overshadowed the simultaneous observation that pancreatic extracts imposed a brief hyperglycemic response prior to the decrease in blood glucose (Banting and Best, 1922; Rosenfeld, 2002). These hyperglycemic properties of pancreatic extracts were first misjudged as an unimportant byproduct of the extraction method. However, in 1922, C.P. Kimball and John R. Murlin identified a pancreatic fraction that potently increased blood glucose when injected in rabbits and dogs (Kimball and Murlin, 1923). Their study concluded that the pancreas produces a specific factor that opposes the hypoglycemic effect of insulin and therefore named the substance "the glucose agonist," or glucagon (Kimball and Murlin, 1923). Prior to the purification of glucagon, research demonstrated that the

hyperglycemic effect of this unidentified substance was partly mediated through a direct glycogenolytic effect on the liver (Burger and Brandt, 1935; Burger and Kramer, 1929). By the mid 1950s, scientists at Eli Lilly successfully crystallized and determined the amino acid sequence of glucagon (Staub et al., 1953; Bromer et al., 1957). These breakthroughs enabled the development of the first radioimmunoassay used subsequently to detect glucagon (Unger et al., 1959) and the medicinal use of glucagon in the treatment severe hypoglycemia (Arieff et al., 1960; Esquibel et al., 1956; MacCuish et al., 1970). The exact place where glucagon was produced in the islets was unknown until 1962, when the previously identified A cells (Lane, 1907), renamed alpha cells, were shown to be the site of glucagon production (Baum 1962; Mikami and Ono, 1962).

Glucagon is the primary secretory product of alpha cells and is the principle measure of alpha cell function. The major physiological stimuli for glucagon release from alpha cells are mixed nutrient or protein ingestion, activation of the autonomic nervous system (ANS), and hypoglycemia (Dunning and Gerich, 2007; Gromada et al., 2007). As suggested by the diversity of these regulatory factors, control of secretion at the level of the islet is complex and multilayered. Interestingly, much of the secretory machinery found in beta cells is also found in alpha cells, including membrane surface receptors, KATP ion channels, and exocytotic proteins (Cheng-Xue et al., 2013; Gromada 2004). Given that alpha and beta cells have opposite secretory responses, these findings were initially difficult to explain (Figure 1-8). However, studies have shown that differences in resting electrical characteristics and ion channel function downstream of KATP channel closure can explain much of the reciprocal pattern of these two hormones and different glycemic conditions (MacDonald et al., 2002; Ramracheya et al., 2010; Rorsman and

Braun, 2012). For example, the glucose transporter Glut2 (encoded by *Slc2a2*), which has an affinity for glucose in the physiological range, is largely absent from alpha cells, whereas the higher-affinity glucose transporter Glut1 (encoded by *Slc2a1*) is present in both mouse (Heimberg et al., 1995; Benner et al., 2014) and human (Blodgett et al., 2015) alpha cells. Thus, the differential expression of GLUT1 and GLUT2 in alpha versus beta cells might affect the range of glucose detection and, conceivably, downstream cellular responses (Gromada, 2018).

In addition to glucose, several paracrine, hormonal, and neuronal signals fine-tune glucagon secretion under different physiological conditions (Gromada et al., 2007). It has long been appreciated that amino acids are important regulators of alpha cells (Holst et al., 2017; Longuet et al., 2013; Solloway et al., 2015; Dean et al., 2017); protein meals or infusions of amino acids stimulate glucagon release, and arginine is potent glucagon secretagogues used in research studies. Furthermore, several studies have provided evidence that insulin suppresses glucagon secretion (Taborsky et al., 2002; Maruyama et al., 1985; Kawamori and Kulkarni, 2009). This effect could be mediated directly via insulin receptors on the alpha cells or indirectly via increased somatostatin secretion from neighboring delta cells (Briant et al., 2018). Lastly, Ishihara and colleagues showed that in the perfused rat pancreas, glucose-stimulated release of zinc ions from beta cells contributed to the inhibition of glucagon secretion (Ishihara et al., 2003). However, examination of mouse islets found that zinc ions did not have an effect on glucagon secretion (Ravier and Rutter, 2005).

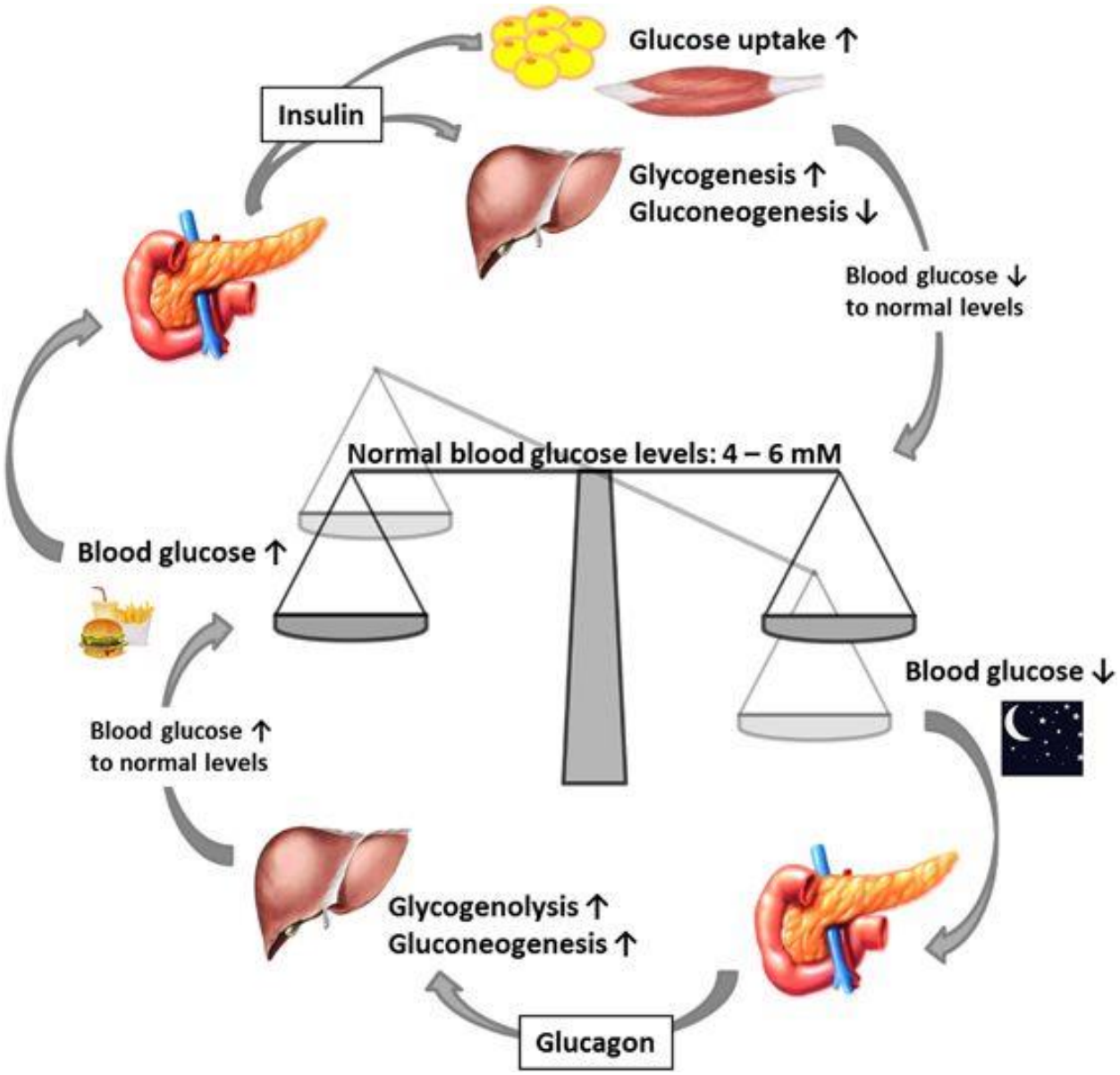


Figure 1-7. Maintenance of blood glucose levels by glucagon and insulin. When blood glucose levels are low, the pancreas secretes glucagon, which increases endogenous blood glucose levels through glycogenolysis. After a meal, when exogenous blood glucose levels are high, insulin is released to trigger glucose uptake into insulin-dependent muscle and adipose tissues as well as to promote glycogenesis. Adapted from Roder et al., 2016.

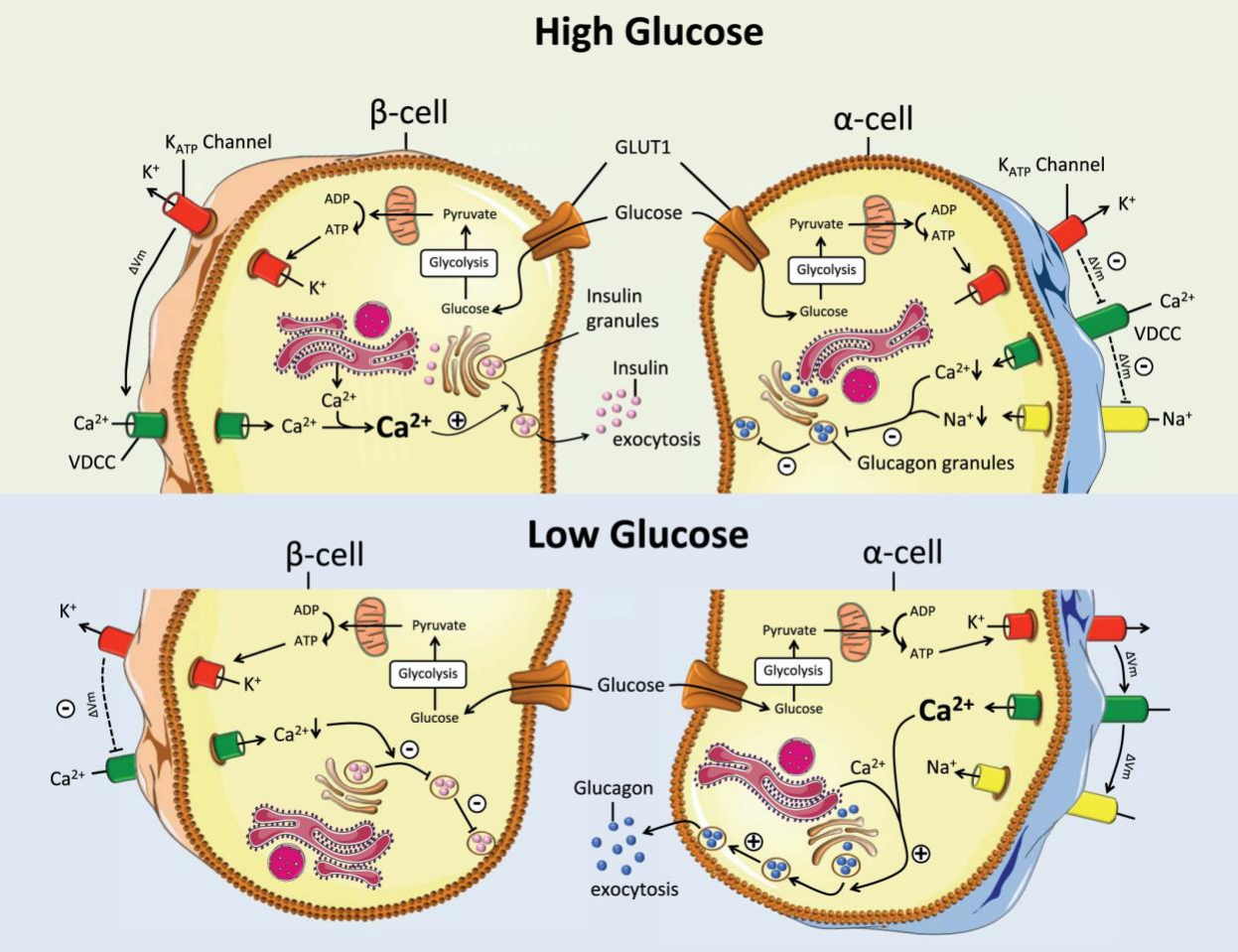


Figure 1-8. Schematic on the glucose-dependent regulation of glucagon and insulin secretion. The cellular machinery controlling the secretion of glucagon in alpha cells is remarkably similar to that regulating the secretion of insulin in beta cells, yet their secretion is reciprocally regulated by blood glucose concentration. In contrast to the secretion of insulin, the release of glucagon is stimulated under conditions of hypoglycemia and subsequently decreases when blood glucose increases. In line with glucose being a major determinant of glucagon (and insulin) secretion, the islets of Langerhans are well vascularized to ensure rapid glucose sensing. Adapted from Muller et al., 2017.

III. Islet dysfunction and diabetes

Health burden

Diabetes represents a complex set of diseases with genetic, immunological, and environmental etiologies. Of the three major types of diabetes, type 2 diabetes mellitus (T2DM) is far more common (accounting for more than 90% of all cases) than either type 1 diabetes mellitus (T1DM) or gestational diabetes; however, T1DM is the most common form of diabetes in children under 15 and over 500,000 children are currently living with this condition worldwide (Katsarou et al., 2014). According to the International Diabetes Federation, 8.8% of the adult population worldwide has diabetes (IDF Diabetes Atlas, 2017). In recent years, there has been an unprecedented increase in diabetes prevalence in the U.S. and worldwide. Sixty years ago, the National Center for Health Statistics (NCHS) began conducting the National Health Interview Survey (NHIS) to gather information on the health of the U.S. population. In 1958, 1.58 million Americans had diabetes, corresponding to the 0.93% of the total population. As of 2015, 30.3 million Americans, or 9.4% of the population, have diabetes (CDC, National Diabetes Statistics Report, 2017). Strikingly, another 84.1 million Americans exhibit impaired glucose tolerance, or prediabetes, a condition that if not treated often leads to diabetes within five years (CDC, National Diabetes Statistics Report, 2017). This rapidly increasing prevalence of diabetes is thought to be due to environmental factors, such as increased availability of food and decreased physical activity, acting on genetically susceptible individuals (Sladek et al., 2007). Diabetes is also currently the seventh leading cause of death in the U.S., corresponding to 76,488 deaths annually (CDC, National Diabetes Statistics Report, 2017). This alarming trend, coupled with limited treatment options and severe comorbidities, highlights the urgent need to address gaps in our understanding diabetes etiology.

Pathophysiology of islet dysfunction

Although the exact pathophysiology differs for each diabetes subtype, the disease generally results from the failure of pancreatic beta cells to meet the insulin demands required for glucose homeostasis. T2DM is a multifactorial disease involving genetic and environmental factors. The disease is characterized by dysregulation of macronutrient metabolism, and results from progressively impaired insulin secretion by pancreatic beta cells, resulting from impaired function, increased cell death, or loss of cell identity (reviewed in DeFronzo, 2015). This beta cell impairment is often in the background of pre-existing insulin resistance in skeletal muscle, liver and adipose tissue. Specifically, beta cell resistance GLP-1 contributes to progressive beta cell dysfunction, while increased glucagon levels and enhanced hepatic sensitivity to glucagon contribute to excessive glucose production by the liver (Magnusson et al., 1992; DeFronzo et al., 2009; DeFronzo et al., 2010). Insulin resistance in adipocytes results in accelerated lipolysis and increased plasma free fatty acid (FFA) levels, both of which aggravate the insulin resistance in muscle and the liver and contribute to beta cell failure (Groop et al., 1989; Guilherme et al., 2008). Increased renal glucose reabsorption by the sodium/glucose co-transporter 2 (SGLT2) and the increased threshold for glucose spillage in the urine contribute to the maintenance of hyperglycemia (Gerich et al., 2001; DeFronzo et al., 2013). Resistance to the appetite-suppressive effects of insulin, leptin, GLP-1, amylin, and PYY, as well as low brain dopamine and increased brain serotonin levels contribute to weight gain, which exacerbates the underlying resistance (Blazquez et al., 2014; Kleinridders et al., 2014; Samuel and Shulman, 2012).

T1DM is a chronic autoimmune disease characterized by persistent hyperglycemia due to insulin deficiency that occurs following destruction of islet beta cells (The SEARCH Study group, 2004;

Gepts, 1965; Eisenbarth, 1986; Atkinson et al., 2014). For the vast majority of patients (70–90%), the loss of beta cells is the consequence T1DM-related autoimmunity, concomitant with the formation of T1DM-associated autoantibodies (American Diabetes Association, 2015). Similar to T2DM, the incidence of T1DM is increasing worldwide and it is estimated that nearly 90,000 children are diagnosed each year (Diaz-Valencia et al., 2015).

The long-standing view of T1DM pathogenesis was that autoimmune beta cell destruction resulted in complete loss of pancreatic insulin secretion. The improved sensitivity of C-peptide detection as well as studies using pancreatic specimens have recently led to the realization that many individuals with T1DM have insulin-secreting cells, even 50 years after diagnosis (Keenan et al., 2010; Oram et al., 2015). Recent work by Brissova and colleagues defined the functional and molecular properties of T1DM islets (Brissova et al., 2018). Surprisingly, they found that remaining islet beta cells maintained several features of regulated insulin secretion. In contrast, islet alpha cells showed severely impaired glucagon secretion, corresponding to decreased levels of essential alpha cell transcription factors (Brissova et al., 2018). These findings highlight the complexity of T1DM pathogenesis and suggest alpha cells play an important role in diabetic hyperglycemia (Brissova et al., 2018).

Current Treatments

The ultimate goal of diabetes care is to minimize fluctuations in blood glucose levels, which will prevent or minimize acute and long-term microvascular complications, such as retinopathy, nephropathy and neuropathy, and macrovascular complications, such as a heart attack and stroke (Pozzilli et al., 2010; Stratton et al., 2000; Holman et al., 2008). The major treatment option for

T1DM is exogenous insulin replacement. More recently, insulin pumps and continuous glucose monitors have made substantial improvements in T1DM care. For T2DM, tight blood glucose control can involve lifestyle interventions (diet and exercise) and/or different kinds of pharmacological agents (Jeon et al., 2007; Sigal et al., 2006). The most common pharmacological agents for T2DM include: suppressors of hepatic glucose production (Metformin), enhancers of insulin sensitivity (thiazolidinediones (TZDs)), and enhancers of insulin secretion (sulfonylureas). Although these agents are the conventional line of treatment, success of these medications varies between patients, and their effectiveness in some cases is only temporary. Furthermore, many of these drugs have several associated risks, including fluid retention, increased cardiovascular risk, lactic acidosis, and dangerous hypoglycemia (Inzucchi et al., 2015).

An alternative treatment for patients with diabetes is the transplantation of human cadaveric islets. This procedure, termed the Edmonton protocol, typically results in better glycemic control, can render patients insulin independent for prolonged periods of time, and improves overall quality of life (Lacy and Scharp, 1986; Mullen et al., 1977; Bellin et al., 2012; Shapiro et al., 2000; Posselt et al, 2010; Barton et al, 2012). However, the severe shortage of cadaveric organ donors and requirement for lifelong immunosuppression impedes the use of islet transplantation as a readily available treatment for people with diabetes (Russ et al., 2015). Consequently, numerous research efforts have focused on identifying alternative sources of surrogate glucose-responsive insulin-producing cells (Zhou and Melton, 2018; Kieffer, 2016; Efrat and Russ, 2012; Hebrok, 2012; Nostro and Keller, 2012; Bouwens et al, 2013; Pagliuca and Melton, 2013). In this next section, I will highlight what is known about the three major strategies for generating

new beta cells: pluripotent stem cell differentiation, reprogramming from other cell types, and induction of replication in existing beta cells (Figure 1-9).

Beta cell replacement therapy

A major advance for the field of beta cell replacement therapy was the identification of pluripotent human ESCs (hESCs) that are capable of generating tissues from all three developmental germ layers (Thomson et al., 1998). In the decade following this discovery, the Yamanaka factors were identified and shown to induce reprogramming of murine fibroblasts into pluripotent stem cells (iPSCs) (Takahashi and Yamanaka, 2006; Wernig et al., 2007; Yu et al., 2007) and soon thereafter, iPSCs were also engineered from human cells (Lowry and Plath, 2008; Nakagawa et al., 2008; Takahashi et al., 2007; Yu et al., 2007). These studies opened the door for research aiming to create fully functional beta cells from pluripotent cells. Two groups were among the first to report differentiation of hESCs towards pancreatic progenitor cells and immature islet cells (D'Amour et al., 2006; Kroon et al., 2008). Importantly, these differentiation protocols were informed by decades of research on pancreas development that mapped out the key steps and critical signaling events required for the formation of mature islets. More recently, several groups have devised enhanced differentiation protocols, which yield cellular clusters with remarkable morphological and functional resemblance to pancreatic islets (Pagliuca et al., 2014; Rezania et al., 2014; Szot et al., 2015; Zhu et al., 2016) (Figure 1-9). Furthermore, transplantation of these clusters into mice led to further functional maturation in vivo and robust rescue of experimental diabetes in mouse models (Pagliuca et al., 2014; Rezania et al., 2014; Szot et al., 2015; Zhu et al., 2016). While these protocols represent a positive step forward in terms of therapeutic development, there are several risks that still need to be assessed, including

1) inadequate control of post- prandial glucose excursions; (2) formation of teratoma, pancreatic adenocarcinoma, or hormone-secreting tumor; (3) hyper-functioning or constitutive functioning of the graft of insulin or other hormones resulting in hypoglycemia; (4) provocation of systemic autoimmunity; and (5) HLA sensitization (Kushner et al., 2014).

Given the complexity of hESC to beta cell differentiation protocols, there is a great interest in converting cell types more closely related to beta cells, such as other islet cell types (Figure 1-9). In the pancreas of adult mice, following near-total beta cell ablation, 1–2% of alpha cells and delta cells spontaneously expressed insulin, leading to correction of diabetic hyperglycemia (Chera et al., 2014; Thorel et al., 2010). A separate study showed that inhibition of the transcription factor Arx, the dimethyltransferase Dnmt1, or forced expression of Pax4, specifically in alpha cells caused transdifferentiation into insulin positive cells, irrespective of beta cell loss (Courtney et al., 2013; Chakravarthy et al., 2017). Additionally, a study from Matsuoka and colleagues showed that forced expression of MafA in pancreatic alpha cells enabled Pdx1-dependent transdifferentiation into beta cells (Matsuoka et al., 2017). Interesting, a recent study showed that a unique population of insulin-producing cells at the periphery of the islets could be an intermediary in the transition from alpha cells to beta cells (van der Meulen et al., 2017). In addition to non-beta islet cells, several studies have explored non-islet cell types as well as extra pancreatic tissues as sources for new beta cells. A combinatorial screening strategy showed that three developmental regulators of beta cells, Ngn3, Pdx1, and MafA, could efficiently convert pancreatic acinar cells into beta-like cells after delivery into the adult mouse pancreas using adenoviral vectors (Zhou et al., 2008). More recently, the same group reassessed the long-term stability of induced insulin+ cells from acinar cells in vivo (Li et al., 2014). This

study showed that by using a newly optimized induction protocol using a polycistronic construct containing Pdx1, Ngn3, and MafA, large numbers of beta-like cells aggregated to islet-like clusters and persisted throughout the observation period up to 13 months (Li et al., 2014). Another group showed that gastrointestinal epithelial cells could also be converted into beta-like cells (Chen et al., 2014). Furthermore, conditional deletion of Foxo1 from Ngn3+ intestinal endocrine progenitors also led to the formation of insulin-producing cells in the gut (Talchai et al., 2012). Other examples of reprogramming mouse cells include: cytokine-mediated conversion of acinar cells to insulin-expressing cells (Baeyens et al., 2014), conversion of duct cells to insulin-expressing cells by FBW7 deletion (Sancho et al., 2014), conversion of Sox9+ ductal cells during development (Kopp et al., 2011), and conversion of hepatocytes to insulin-producing cells by TGIF2 (Cerda-Esteban et al., 2017). Despite numerous proof-of-concept demonstrations of beta cell reprogramming in animal models, efforts to reprogram human cells have been less successful. While several studies have suggested that human alpha cells can be reprogrammed to become beta cells (Ben-Othman et al., 2017; Xiao et al., 2018), in the absence of lineage tracing tools, direct evidence is still lacking. At this point, the main challenge in translating the reprograming approach into the clinic is to define reliable methods for efficient production of human beta-like cells that can develop stable and functional transplants.

Stimulating beta cell proliferation is a simple and intuitive solution to replenishing beta cell mass (Dor et al., 2004; Nir et al., 2007; Zhou and Melton, 2018) (Figure 1-9). In fact, a large number of growth factors and mitogenic agents have been shown to promote beta cell proliferation in animal models, including parathyroid hormone-related protein, hepatocyte growth factor, GLP-1, insulin-like growth factors, gastrin, epidermal growth factors, platelet-derived growth factor, and

adenosine kinase inhibitors (Mezza and Kulkarni, 2014; Wang et al., 2017; Saunders and Powers, 2016; Wang et al., 2015; Andersson et al., 2012; Schulz et al., 2016; Annes et al., 2012). Unfortunately, these agents have generally failed to promote significant proliferation of human beta cells. Evidence suggests that substantial proliferation of human beta cells appears to occur naturally only in the first year of life (Kassem et al., 2000; Meier et al., 2008; Gregg et al., 2012). There is longstanding evidence that insulin and glucose, both of which are elevated in obesity or insulin resistance, may directly stimulate beta cell proliferation (Kulkarni, 2005; Dadon et al., 2012; Stamateris et al., 2016). But it remains unclear whether these are the key signals that drive islet hyperplasia. A potentially important advance has come from high-throughput compound screens that identified inhibitors of dual specificity tyrosine-phosphorylation-regulated kinase 1A (DYRK1A) as reagents that can potently stimulate proliferation of cultured human beta cells in vitro and transplanted human beta cells in vivo (Wang et al., 2015; Dirice et al., 2016; Shen et al., 2015). In addition, other pathways involved in human beta ell proliferation, such as calcineurin and SerpinB1, have been identified (El Ouaamari et al., 2017; Dai et al., 2017). To advance these reagents into clinics will require controlling the cell type specificity, targeting the intervention to islets, and ensuring that reagents that do not raise the problem of tumor formation. Ultimately, a true cure for diabetes will require the elucidation of the molecular and genetic etiology of the disease.

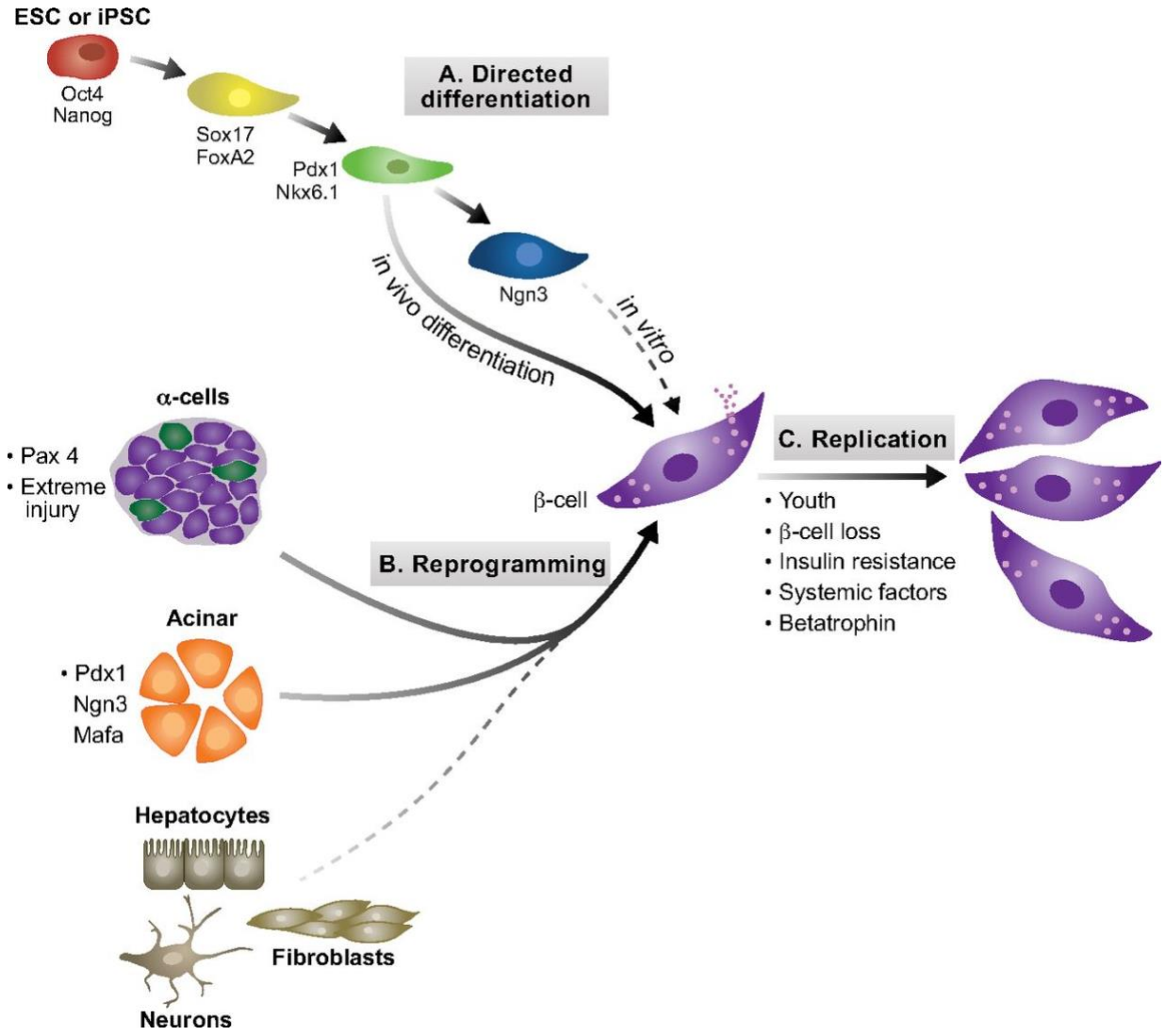


Figure 1-9. Strategies to generate new beta cells. (A) Directed differentiation using growth factors and small molecules can direct a pluripotent stem cell (red) through the stages of pancreatic differentiation in a manner that mimics normal development. Currently, functional beta cells can only be differentiated through an in vivo transplantation step, but deriving a bona fide beta cell fully in vitro (dashed line) is a major goal. A subset of important genes expressed at each stage is listed. (B) Reprogramming of terminally differentiated cell types, such as acinar or alpha cells, can be used to generate beta cells in vivo, using the overexpression or injury strategies listed. Reprogramming other mature cell types, such as hepatocytes, fibroblasts or

neurons, in vitro into beta cells (dashed line) remains to be achieved. (C) Inducing the replication of existing beta cells is the primary strategy for generating new endogenous beta cells. Replication may be recapitulated in vitro or induced in vivo with new small molecules or proteins based on the strategies listed. Adapted from Paglucia and Melton, 2013.

Genetic factors contributing to diabetes

Decades of mouse research have yielded major advances in our understanding of factors influencing beta cell development, identity, and dysfunction. Specifically, these studies have revealed networks of transcriptional and chromatin regulators that drive beta cell lineage development (Oliver-Krasinski and Stoffers, 2008; Arda et al., 2013; van Arensbergen et al., 2010; Bramswig et al., 2013; Xu et al., 2011), provide barriers against transdifferentiation or loss of beta cell identity (Talchai et al., 2012; Dhawan et al., 2011; Puri et al., 2013; Taylor et al., 2013; Gao et al., 2014; Papizan et al., 2011; Swisa et al., 2017; Ediger et al., 2017; Collombat et al., 2009; Gutierrez et al., 2017), and are misregulated in T1DM and T2DM (Lu et al., 2018; Pedersen et al., 2017; Wang et al., 2015; Hou et al., 2017; Carrero et al., 2013). Additionally, advances in genome wide association studies (GWAS) have identified several genetic drivers of monogenic syndromes of beta cell dysfunction, as well as 113 distinct T2DM susceptibility loci and 60 loci associated with an increased risk of developing T1DM (McCarthy 2010; McCarthy 2016; Scott et al., 2017; Zhao et al., 2017; Cropano et al., 2017; Soccio et al., 2015; Roman et al., 2017; Rusu et al., 2017; Maurano et al., 2012).

Strikingly, these studies discovered that most T2DM and T1DM susceptibility loci fall outside of coding regions, which suggests a role for noncoding elements in the development of disease. Exceptions are a few variants in exons, which influence the function of the gene, such as *SLC30A8* (Flannick et al., 2014), *KCNJ11,* and *GCKR* (Sladek et al., 2007; Saxena et al., 2007). Several studies have demonstrated that many causal variants of diabetes are significantly enriched in regions containing islet enhancers, promoters, and transcription factor binding sites; however, associations with these regulatory regions cannot explain all diabetes susceptibility.

This highlights our incomplete understanding of the islet regulome and suggests the need for detailed functional analyses of noncoding genes to precisely determine their contribution to diabetes susceptibility and disease progression.

IV. Noncoding RNAs

Regulatory noncoding RNAs

Advances in RNA sequencing (RNA-seq) technologies have revealed that mammalian genomes encode tens of thousands of RNA transcripts that have similar features to mRNAs, yet are not translated into proteins. While the percentage of non-protein coding RNAs that are functional is still unknown, detailed characterization of many of these transcripts has challenged the idea that the central role for RNA in a cell is to give rise to proteins. Instead, these RNA transcripts make up a class of molecules called noncoding RNAs (ncRNAs) that function either as "housekeeping" ncRNAs, such as transfer RNAs (tRNAs) and ribosomal RNAs (rRNAs), that are expressed ubiquitously and required for protein synthesis, or as "regulatory" ncRNAs that control gene expression. Regulatory ncRNAs are further subdivided according to size: 1) short ncRNAs (~20 nucleotides), such as micro RNAs (miRNAs) and small interfering RNAs (siRNAs); and 2) long noncoding RNAs (lncRNAs) that are defined as transcripts longer than 200 nucleotides (Esguerra and Eliasson 2014).

MicroRNAs in beta cell development and function

MicroRNAs are the most well-characterized class of regulatory ncRNAs. MiRNAs are generally believed to negatively regulate gene expression through mRNA cleavage that is mediated by the Argonaute (Ago) family of proteins as part of the RNA-induced silencing complex (RISC) (Mendell and Olson, 2012). The finding that miRNAs only require short "seed" sequences in the 5' end of their transcript to direct Ago to target mRNAs implies that one miRNA can influence the expression of several target genes (Bartell 2009). It is therefore not surprising that miRNAs play essential regulatory roles in diverse cellular processes, including proliferation,

organogenesis, hormone secretion, and apoptosis (Esguerra and Eliasson 2014). Studies have also shown that the misexpression of miRNAs disrupts gene regulatory networks and contributes to disease states, such as cancer and diabetes (Mendell and Olson, 2012). In fact, several groups have identified essential roles for miRNAs in beta cell development, function, and disease.

The requirement of miRNAs in beta cell development was initially demonstrated by removal of Dicer function during different stages of pancreas development (Figure 1-10). Given that Dicer is required for the formation of mature miRNAs, loss of Dicer ablates all miRNA function within a cell (Bernstein et al., 2003). The pan-pancreatic loss of Dicer using the Pdx1:Cre allele resulted in severe pancreatic agenesis and neonatal death (Lynn et al., 2007). The mutant mice also displayed a dramatic loss of all endocrine cell types, which was attributed to a reduction in the number of Ngn3+ endocrine progenitor cells (Lynn et al., 2007). In a more recent study, Ngn3:Cre was used to ablate Dicer function specifically in the endocrine progenitor population. The mutant mice had normal embryonic pancreas development; however, shortly after birth the mice developed hyperglycemia due to progressive loss of beta cells (Kanji et al., 2013). Surprisingly, the authors found that loss of miRNA function in endocrine progenitor cells caused upregulation of neuronal genes normally repressed by RE1-silencing transcription factor (REST) (Kanji et al., 2013). Thus, it appears that miRNAs maintain islet-cell identity by inhibiting expression of REST, thereby restricting expression of neuronal genes.

While disruption of Dicer function shows that miRNAs are generally required for proper pancreas development and islet cell specification, individual miRNAs have also been implicated in beta cell specification (Figure 1-10). Kredo-Russo et al. showed that miRNA-7 (miR-7) is

expressed during endocrine pancreas development where it directly targets and controls expression levels of the essential pancreatic transcription factor, Pax6 (Kredo-Russo et al., 2012). Additionally, the miR-30 family functions during the epithelial-to-mesenchymal transition by inhibiting mesenchymal mRNAs, such as Vimentin and Snail1, to favor pancreatic epithelial development (Joglekar et al., 2009). Taken together, these studies define a requirement for miRNAs in several layers of beta cell development.

Following beta cell specification, complex networks of regulatory factors, including miRNAs, are needed to maintain beta cell function. Two groups have examined the role of Dicer specifically in the developing beta cells by using the Rat Insulin Promoter (RIP):Cre (Kalis et al., 2011; Mandelbaum et al., 2012). Unlike mice with Dicer deleted throughout the whole pancreas, the beta cell specific Dicer mutant mice survived postnatally, but developed a diabetic phenotype around 8-weeks of age due to an overall decrease of beta cell number and reduced insulin production and secretion (Kalis et al., 2011; Mandelbaum et al., 2012). Ultrastructural analysis demonstrated that the Dicer mutant beta cells had 50% fewer insulin granules than control beta cells (Kalis et al., 2011). Intriguingly, mutant islets also had a dramatic reduction in insulin transcript levels, suggestive of a defect in insulin gene transcription (Kalis et al., 2011). Global miRNA function also appears to be required for maintenance of beta cell identity. Ablation of Dicer in adult beta cells (Melkman-Zehav et al., 2011) using tamoxifen inducible RIP:CreER;Dicer$^{fl/fl}$ mice, resulted in hyperglycemia, glucose intolerance and a drastic reduction in beta cell number (Melkman-Zehav et al., 2011). The authors concluded that the loss of miRNA function caused the upregulation of several transcriptional repressors, including Sox6 and Bhlhe22, which in turn triggered decreased insulin expression (Melkman-Zehav et al., 2011).

Individual miRNAs have also been shown to be essential for many aspects of beta cell function (Figure 1-10). The most highly expressed miRNA in mouse and human islets, miRNA-375, was first identified in a murine pancreatic beta cell line (MIN6) where it was shown to negatively regulate glucose-stimulated insulin secretion (Poy et al., 2004). Analyses in mice showed that genetic ablation of miR-375 resulted in hyperglycemia due to reduced beta cell mass (Poy et al., 2009). This defect was attributed to a significant reduction in beta cell proliferation due to the upregulation of several genes that negatively regulate cell growth (Poy et al., 2009). Several miRNAs including miR-24, miR-26, miR-148, and miR-182 were also shown in vitro to negatively regulate insulin expression (Melkman-Zehavi et al., 2012). However, the relatively minimal effect seen with individual miRNA knockdown suggests that a combination of multiple miRNAs maintain insulin expression (Melkman-Zehavi et al., 2012). Interestingly, the evolutionarily conserved miRNA, miR-7a, was shown to be a negative regulator of both insulin secretion and beta cell proliferation through inhibition of mTOR signaling proteins, indicating that a single miRNA can also regulate multiple layers of beta cell function (Latreille et al., 2014; Wang et al., 2013).

Consistent with recent implications that miRNAs maintain tissue homeostasis during a cellular response to stress, several studies have implicated miRNAs in beta cell stress and diabetes (Mendell and Olson, 2012). Recently, Latreille and colleagues elucidated the relationship between miR-7a expression levels and two well-known diabetes mouse models (Latreille et al., 2014). Specifically, they determined that mice with a beta cell specific miR-7a deletion had a similar phenotype to genetically obese (*ob/ob*) mice; mice had enhanced pancreatic beta cell function due to compensatory mechanisms that allowed them overcome insulin resistance

(Latreille et al., 2014). This finding was supported by a corresponding 50% reduction in miR-7a transcript levels in compensating islets from *ob/ob* mice (Latreille et al., 2014). Remarkably, they found the opposite defect when they overexpressed miR-7a in vivo: hyperglycemia, reduced plasma insulin levels, and impaired glucose-stimulated insulin secretion (Latreille et al., 2014). Overexpression of miR-7a also caused decreased *Ins1 and Ins2* expression, with a correlating decline in expression of several mature beta cell markers (Latreille et al., 2014). Of note, these phenotypes are strikingly similar to the metabolic phenotype of diabetic *db/db* mice, which develop hyperglycemia and have reduced plasma insulin levels over time due to beta cell dysfunction (Latreille et al., 2014). Mechanistically, miR-7a was shown to directly regulate genes that control late stages of insulin granule fusion with the plasma membrane, indicating a direct role for miRNAs in regulating beta cell function during disease (Latreille et al., 2014). Similarly, the finding that miR-375 is upregulated in islets isolated from *ob/ob* mice, combined with an established role for miR-375 in beta cell proliferation, suggests a mechanism whereby miR-375 enables beta cell proliferation to compensate for metabolic stress (Poy et al., 2009). Tattikota et al. further elucidated a mechanistic link between miR-375, Ago2, and another miRNA, miR-184 (Tattikota et al., 2014). This study determined that the onset of insulin resistance caused the silencing of miR-184, which then released its constraint on Ago2, a major component of the RNA-induced silencing complex (Tattikota et al., 2014). Ago2 is then able to orchestrate the suppressive function miR-375 thus enhancing beta cell proliferation to accommodate the physiological demand for insulin (Tattikota et al., 2014). Taken together, these studies have shown that miRNAs work to maintain beta cell function during metabolic stress, and aberrant miRNA expression can be a marker for beta cell dysfunction (Figure 1-10).

Characterizing miRNAs that regulate beta cell function may also reveal novel methods to treat diabetes. For example, the Let-7 miRNA family was shown to negatively regulate glucose-stimulated insulin secretion (Zhu et al., 2011). A recent study sought to determine if global Let-7 knockdown was sufficient to prevent or rescue impaired glucose intolerance in mice (Joglekar et al., 2009). Researchers put mice on a high fat diet for 10 weeks and then initiated weekly injections with a locked nucleic acid (LNA) modified anti-miR that ablated Let-7 function (Joglekar et al., 2009). After confirming strong reduction of Let-7 transcript in several tissues, researchers found that Let-7 knockdown was sufficient to prevent and treat impaired glucose tolerance brought on by a high fat diet (Joglekar et al., 2009). Overall, miRNAs have been shown to be critical regulators of beta cell development and function. Harnessing miRNA therapeutic capabilities will require a more comprehensive understanding of their mechanism of action in the pancreas.

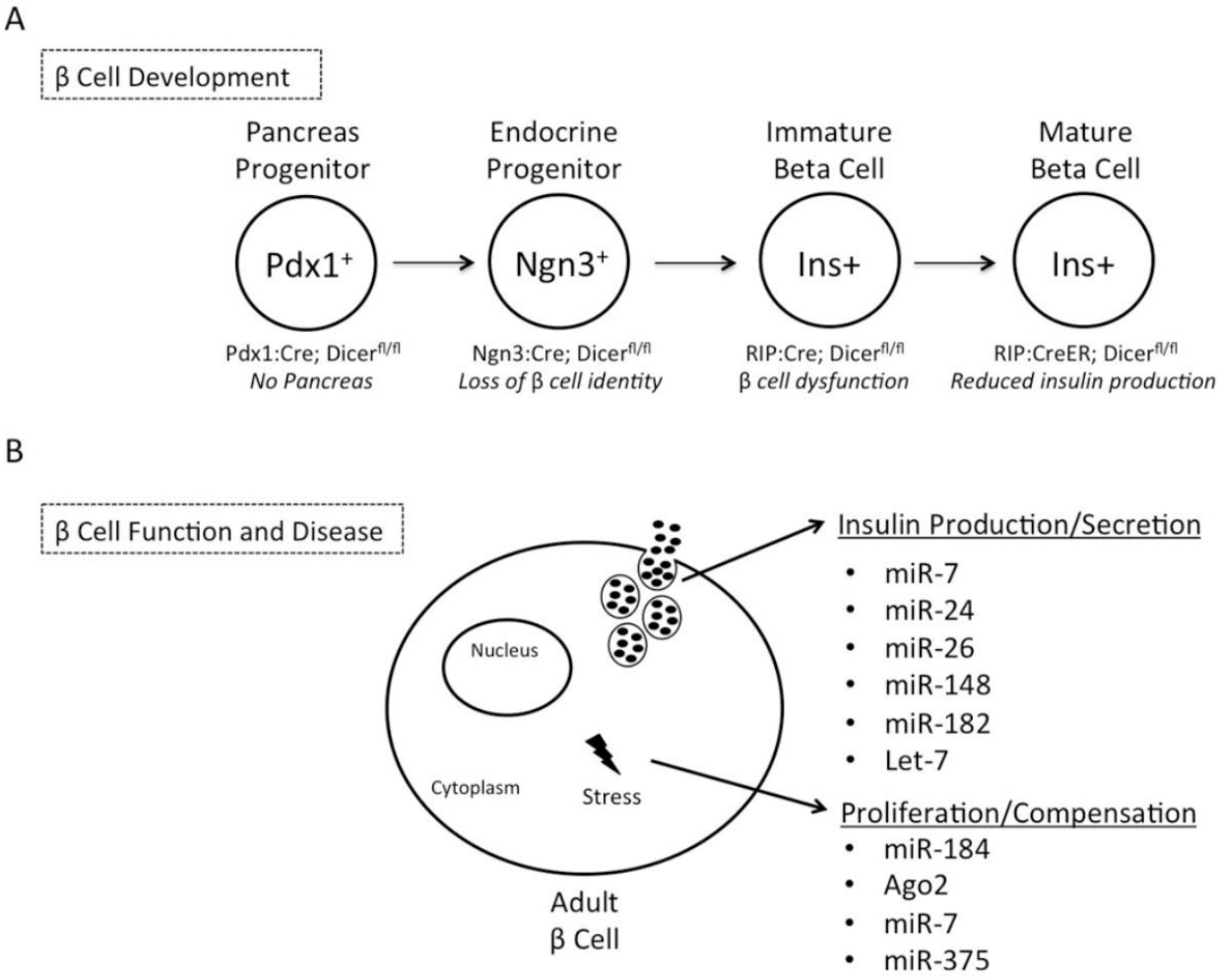


Figure 1-10. The role of microRNAs in beta cell development, function, and disease. (A) A pictorial summary of the studies that identified a role for global miRNA function during beta cell development through temporal ablation of Dicer in the pancreas in vivo. Each major stage of pancreas development is represented along with the genotype of conditional Dicer ablation and corresponding phenotype. (B) Summary of a subset of individual miRNAs that regulate beta cell function. The indicated miRNAs regulate glucose-stimulated insulin secretion and/or play a role in compensatory mechanisms during beta cell stress. Adapted from Singer, Arnes, and Sussel, 2015.

Long noncoding RNAs

While the functional mechanisms of short regulatory ncRNAs, predominantly microRNAs (miRNAs), have been described in detail, the most abundant and functionally enigmatic regulatory ncRNAs are called long noncoding RNAs (lncRNAs) that are loosely defined as RNAs larger than 200 nucleotides (nt) and do not encode for protein (Derrien et al., 2012; Batista and Chang, 2013; Rinn and Chang, 2012). While using a definition based strictly on size is somewhat arbitrary, this definition is useful both bioinformatically (the need for an effective size gap to distinguish lncRNAs from the ~20 nt short ncRNAs) and technically (200 nt is approximately the retention size of most silica-based columns used in standard RNA extraction kits) (Cabili et al., 2011). Although the 200 nt size cutoff has simplified identification of lncRNAs, this rather broad classification means several features of lncRNAs, including abundance, cellular localization, stability, conservation, and function are inherently heterogeneous (Ponting et al., 2009; Wilusz et al., 2009; Guttman and Rinn, 2012). Although this represents one of the major challenges of lncRNA biology, it also highlights the untapped potential of lncRNAs to provide a novel layer of gene regulation that influences islet physiology and pathophysiology.

Major efforts to unify the lncRNA field by clarifying nomenclature and identifying common attributes of lncRNAs have yielded several important conclusions about their shared characteristics (Guttman and Rinn, 2012). In general, despite a lack of translatable open reading frames, most lncRNAs are biochemically indistinguishable from their mRNA counter- parts: they are transcribed by RNA polymerase II (Pol II) from genetic loci that contain classic active chromatin marks at their promoters (H3K4me3) and gene bodies (H3K36me3), and they are 5’

capped, spliced, and polyadenylated (Guttman et al., 2009). Beyond these common characteristics, lncRNAs represent a heterogeneous population of functional RNAs with diverse functions that have yet to be fully explored. As the lncRNA field has grown, several databases, such as NONCODE (Luo et al., 2017), lncRNAdb (Amaral et al., 2011), and LNCipedia (Volders et al., 2013), have emerged to provide researchers with catalogs of empirically identified lncRNAs (Figure 1-11). Additional attempts to categorize the myriad of newly identified lncRNAs have yielded several classifications based on genomic proximity to the nearest protein-coding gene (Mercer et al., 2009). Although it can be useful to have classifications that do not rely on functional annotation, the genomic contexts of lncRNAs do not necessarily provide insights into function. For example, two well-characterized lncRNAs, HOTTIP (HOXA transcript at the distal tip) and HOTAIR (HOX transcript antisense RNA), are both transcribed from the HOXA gene locus; however, HOTTIP regulates a nearby HOXA gene in cis, whereas HOTAIR recruits chromatin-modifying complexes to other chromosomes in trans (Wang et al, 2011; Rinn et al., 2007). Functional classification based on sequence alone has also been difficult because lncRNAs are not under the same evolutionary pressure as proteins to convey a genetic code (Smith and Mattick, 2007). Instead, information from the small number of well-characterized lncRNAs has been distilled to describe four non-mutually exclusive archetypes of lncRNA molecular mechanisms: signals, decoys, guides, and scaffolds (Wang and Chang, 2011). These archetypes emphasize the unique ability of lncRNAs to bind to DNA, RNA, and proteins to regulate all layers of gene expression. As more lncRNAs undergo functional characterization, improved classifications will enable predictive modeling of lncRNA function.

V. Islet long noncoding RNAs

Long noncoding RNAs and diabetes

Although the role of miRNAs in diabetes has been well established (LaPierre and Stoffel, 2017), analyses of lncRNAs in islets have lagged behind their short ncRNA counterparts. However, several recent studies provide evidence that lncRNAs are crucial components of the islet regulome and may have a role in diabetes (Motterle et al., 2016). For example, a genome-wide association study examined T2DM susceptibility loci within unknown genomic associations and found that a significant percentage (> 16%) of T2DM loci contained islet lncRNAs within 150 kb of the reported lead single nucleotide polymorphism (SNP) (Moran et al., 2012), suggesting lncRNAs are essential for normal pancreatic function. Furthermore, misexpression of several lncRNAs has been correlated with diabetes complications, such as diabetic nephropathy and retinopathy (Alvarez et al., 2011; Hanson et al., 2007; Awata et al., 2014). There are also preliminary studies suggesting that circulating lncRNAs, such as *Gas5, MIAT1, and SENCR,* may represent effective molecular biomarkers of diabetes and diabetes-related complications (Carter et al., 2015; de Gonzalo-Calvo et al., 2016). Finally, several recent studies have explored the role of lncRNAs in the peripheral metabolic tissues that contribute to energy homeostasis (reviewed in Giroud and Scheideler, 2017).

In addition to their potential as genetic drivers and/or biomarkers of diabetes and diabetes complications, lncRNAs can be exploited for the treatment of diabetes. For example, although tremendous efforts have been dedicated to generating replacement beta cells for individuals with diabetes (Pagliuca et al., 2014; Russ et al., 2015), human pluripotent stem cell–based beta cell differentiation protocols remain inefficient, and the end product is still functionally and

transcriptionally immature compared with primary human beta cells (reviewed in Millman and Pagliuca, 2017). This is largely due to our incomplete knowledge of in vivo differentiation regulatory pathways, which likely include a role for lncRNAs. Once we gain additional understanding of lncRNA function during the pancreatic endocrine differentiation process, we can incorporate lncRNAs into the in vitro differentiation protocols to optimize the generation of beta cells derived from human pluripotent stem cells.

Inherent characteristics of lncRNAs have also made them attractive candidates for drug targeting, which could be exploited for developing new diabetes therapies. LncRNAs can be targeted by antisense oligonucleotides (ASOs), which is technically simpler than small molecule screening or inhibitory antibody development. Also, as strategies to therapeutically target mRNA molecules date back to the 1970s, developing next-generation technologies to target lncRNAs is quite feasible and easily adapted from traditional mRNA targeting approaches. For example, several currently available biochemical modifications to ASOs, such as a LNA backbone, increase their stability and reduce toxicity (Adams et al., 2017). Unique features of lncRNAs also make them more amenable to targeting, including their highly tissue-specific expression and their relatively lower expression levels as compared with coding mRNAs, which may allow the use of lower doses of targeting molecules, thus alleviating drug toxicity (Derrien et al., 2012; Cabili et al., 2014). A study in human pancreatic islets found that although 9.4% of Ref-seq–annotated protein-coding genes were islet specific, 55% of intergenic lncRNAs were expressed only in islets (Moran et al., 2012). Taken together, these findings suggest that lncRNAs could be easily targeted and their restricted expression profile would result in fewer unwanted pleiotropic phenotypes.

Identification of islet lncRNAs

Although the rapid evolution of the lncRNA field has provided remarkable insight into lncRNA functions in many different organ systems and disease states (Derrien et al., 2012; Cabili et al., 2011; Guttman et al., 2009; Mercer et al., 2009), these studies have also revealed important aspects of lncRNA biology that need be considered to optimize a discovery pipeline (Figure 1-11). Unlike the discovery of novel coding genes, whose functions can be inferred from the functional protein domains they encode, a different set of strategies must be used for lncRNAs. To optimally characterize and exploit lncRNAs for treating diabetes, it will be useful to develop a common set of guidelines to most effectively identify those lncRNAs that are essential for pancreatic function. With this goal in mind, we will discuss studies that have identified mouse and human pancreatic lncRNAs, highlight the major steps of lncRNA discovery pipelines for diabetes-related lncRNAs, and identify the advantages and disadvantages of each approach. The information gained from these pioneering studies will inform the identification of functionally relevant islet-specific lncRNAs that can be exploited to promote the efficient generation or regeneration of beta cells or can be targeted for novel therapies to treat diabetes. The first systematic mapping of lncRNAs in the pancreas was described in 2012; Moran et al performed deep RNA-seq on human islets in combination with chromatin immunoprecipitation sequencing for three epigenetic markers of active transcription: H3K4me3, H3K36me3, and RNA Pol II (Moran et al., 2012). This strategy made it possible to leverage the histone marks to identify 1128 human islet lncRNAs that showed classic marks of active transcription (Moran et al., 2012). More recently, this same group reported a similar, updated set of parameters including the presence of H3K4me3 enrichment in a region +1 kb to -0.5 kb from a transcript's 5' end to identify 2226 human islet lncRNAs (Akerman et al., 2017). Subsequently, Fadista et al used

associations with disease loci to identify potential diabetes-related lncRNAs; genetic variants regulating gene expression, referred to as expression quantitative trait loci (eQTLs), that were implicated in altered glucose metabolism and insulin secretion determined that 33 out of 616 diabetic eQTLs influenced the expression of human islet lncRNAs (Fadista et al., 2014). As the human diabetes field moves beyond exomic sequencing to whole-genome sequence analysis, it is likely that additional lncRNAs will be identified as contributing factors to the etiology of diabetes.

Although using human pancreatic islets as a source of RNA has the advantage of identifying lncRNAs most relevant for human disease (Moran et al., 2012; Akerman et al., 2017; Fadista et al., 2014; Bramswig et al., 2013; Li et al., 2014), cadaveric islets are not only scarce but also often highly heterogeneous and have reduced RNA integrity, which can result in RNA artifacts and con- found the identification of bona fide lncRNA molecules. Furthermore, the options for downstream functional analyses of human lncRNAs are limited. As an alternative, several lncRNA screens have been performed in mouse islets (Ku et al., 2012; Benner et al., 2014; Motterle et al., 2017). This approach benefits from higher-quality starting material, increased sample homogeneity, and the ability to use cell-specific reporters to isolate purified islet cell populations. Rodent models also afford the option to interrogate lncRNA function in vivo. Furthermore, although lncRNAs are often poorly conserved at the nucleotide level between mice and humans, Moran et al reported that 70% of human islet lncRNAs had an orthologous mouse genomic region and that 47% of those orthologous mouse loci produced a corresponding lncRNA transcript (Moran et al., 2012). Currently, there is a paucity of information regarding the minimal amount of nucleotide conservation that would predict conserved function, especially as

the conservation may be at the level of RNA secondary structure (Ulitsky et al., 2011). However, the characterization of increased numbers of conserved orthologous lncRNAs will pave the way to understanding the contribution of lncRNA function to beta cell biology and inform whether lncRNAs that are identified and characterized in mice contribute to human beta cell function and disease pathologies.

Expression patterns of islet lncRNAs

As it remains difficult to use lncRNA nucleotide sequence as a predictor of functional activity, cell-specific and/or regulated expression can be used to predict lncRNA function. Although most transcriptional regulatory proteins are expressed in many different tissues and cell types, expression of functionally important lncRNAs tends to be much more restricted (Cabili et al., 2011). This can present a challenge for the identification of lncRNAs as expression must be assessed in the correct cell type, but it also provides the exciting possibility that lncRNAs confer highly specialized cell-specific regulatory functions. To identify islet cell-type specific lncRNA transcripts, a study used beta cell-specific reporter mice (MIP:GFP) to compare the transcriptomes of MIP:GFP-positive mouse islet cells (beta cells) to MIP:GFP-negative (non-beta cell) islet cells and discovered ~12% of islet lncRNAs were beta cell specific (Ku et al., 2012). Similarly, studies have identified differentially expressed lncRNAs in human alpha versus beta cells (Bramswig et al., 2013) and at different stages of in vitro human pancreas differentiation (Jiang et al., 2015). Given the highly cell-type specific expression of lncRNAs, it is tempting to speculate that these molecules provide the regulatory specificity to gene networks that define cell-specific identities and functions. For example, the lncRNA βlinc1 is expressed exclusively in pancreatic beta cells, where it appears to regulate a beta cell-specific regulatory

program (Arnes et al., 2016). Coexpression and/or cross-regulation of lncRNAs with a nearby protein-coding gene can also provide important functional and regulatory information; the human islet lncRNA *PLUTO* had a highly correlated tissue-specific expression pattern with its neighboring gene, the essential pancreatic transcription factor PDX1, and was further shown to directly regulate PDX1 transcription (Akerman et al., 2017).

Altered expression in pathophysiological conditions could also be a defining characteristic in the identification of functionally important lncRNAs. The lncRNAs *Meg3* and *Tug1* were both shown, in separate studies, to be downregulated in the NOD T1DM mouse model (Yin et al., 2015; You et al., 2016). Similarly, a study in human islets found that two lncRNAs *KCNQ1OT1* and *HI-LNC45* were significantly upregulated or downregulated in T2DM islets, respectively (Moran et al., 2012). Several groups also identified lncRNAs misregulated in islets exposed to altered stimuli or physiological challenges, such as high glucose (Benner et al., 2014), high-fat diet (Motterle et al., 2017), cytokines (Motterle et al., 2015), and pregnancy (Sisino et al., 2017). Interestingly, although exposure of mouse islets to inflammatory cytokines caused upregulation of four lncRNAs (Li et al., 2014), a study on human islets cultured in the presence or absence of cytokines found no differential expression of lncRNAs between the two conditions (Motterle et al., 2015), suggesting these analyses may be influenced by low lncRNA expression levels and detection limits of the assay.

Subcellular localization can yield perhaps the most telling clues about how lncRNAs function mechanistically. Nuclear lncRNAs are more likely to regulate transcription, whereas cytoplasmic lncRNAs more often influence translation or mRNA stability (Wang and Chang, 2011). An

individual lncRNA can also be expressed in both the nucleus and cytoplasm and have a different function in each compartment (Kino et al., 2010). Although subcellular fractionation has traditionally been used to assess cellular localization, RNA fluorescent in situ hybridization (FISH) is a highly sensitive assay that can also provide crucial spatial information about lncRNA localization within the cell. For example, *DEANR1* was shown to directly regulate the transcription factor FOXA2 in differentiated endoderm by using RNA-DNA-FISH experiments showing that DEANR1 localized to two punctae corresponding to the FOXA2 DNA locus (Jiang et al., 2015).

Although expression analyses can be performed using a combination of quantitative reverse transcription PCR (qRT-PCR) and RNA in situ hybridization, in silico discovery of islet lncRNAs is also possible as a result of the comprehensive mapping of mouse alpha-, beta-, and delta-cell transcriptome (DiGruccio et al., 2016). Additionally, global expression information can often be obtained in silico through publically available data sets, such as BodyMap (Li et al., 2017); however, these resources rarely include pancreas and/or islet RNA in their arrays due to challenges associated with pancreatic RNA integrity. Furthermore, standard RNA-seq parameters often lack the coverage needed to identify low abundant lncRNAs, many of which remain unannotated. Another obstacle to lncRNA discovery is the paucity of relevant tissue, given that islets make up only about 5–10% of the entire pancreas. As new molecular technologies with increased detection sensitivities are developed, these challenges will be overcome; however, this will also increase the importance of functional validation for newly identified pancreatic lncRNAs.

Functional Characterization of Islet lncRNAs

With the advancement of high-throughput sequencing techniques, the list of islet-specific lncRNAs is growing exponentially; however, functional characterization is missing for the majority of these lncRNAs. Studies that have experimentally determined the function of an islet lncRNA have used several different approaches. The most straightforward strategy to test the regulatory function of islet lncRNAs is by loss-of-function assays using RNA interference (RNAi). Knockdown (KD) of the lncRNA *βlinc1* in MIN6 cells demonstrated that *βlinc1* is a novel cis regulator of the islet transcription factor Nkx2–2 (Arnes et al., 2016), whereas short hairpin RNA (shRNA) KD of the lncRNA *DEANR1* led to a drastic downregulation of the nearby gene FOXA2 (Jiang et al., 2015). In vitro RNAi is also the optimal technique to identify lncRNAs that function in trans to regulate essential islet genes: *HI-LNC25* was shown to regulate the distant islet transcription factor GLIS3 in EndoC-bH1 cells (Moran et al., 2012), the lncRNA *Meg3* positively regulated Pdx1 and MafA in MIN6 cells (You et al., 2016), and the lncRNA-ROR regulated insulin, Pdx1, and Glut2 in human amniotic epithelial cells differentiated into beta-like cells (Zou et al., 2016). Furthermore, KD of lncRNA *Tug1* in NIT-1 cells was correlated with changes in beta cell function, including decreased glucose-stimulated insulin secretion (Yin et al., 2015). In a more comprehensive screen, Akerman and colleagues functionally interrogated 12 human lncRNAs with lentiviral vectors containing Pol II–transcribed artificial miRNAs (coined amiRNAs) with perfect homology to the target lncRNA to elicit degradation via the RNAi pathway (Akerman et al., 2017). Remarkably, KD of 9 out of 12 lncRNAs in EndoC-bH1 cells elicited significant gene expression changes, and KD of three of those lncRNAs led to impaired insulin secretion (Akerman et al., 2017).

Although siRNAs and shRNAs are well suited to downregulate the expression and/or translation of protein-coding genes, there is concern in the field about the use of cytoplasmic RNAi machinery to KD nuclear lncRNAs (Basset et al., 2014). Modified ASOs, such as LNA GapmeRs, are a valuable alternative for lncRNAs enriched in the nucleus. They are stable high-affinity RNA analogs that readily permeate the nucleus and function by RNase H-dependent degradation of complementary RNA targets (Xing et al., 2014). Gain-of-function experiments, mainly lncRNA overexpression, should also be considered when overexpression more closely mimics an endogenous or diseased state, as was the case for four lncRNAs shown to be upregulated in MIN6 cells exposed to inflammatory cytokines (Motterle et al., 2015).

Although in vitro functional analysis has been informative, there are caveats associated with using immortalized cell lines. For example, several lncRNAs with significant functions in vitro have had no phenotype when knocked out (KO) in mice (Bassett et al., 2014). This discrepancy may be due to improper use of KD tools or because lncRNA KO phenotypes can be subtle and may only appear after physiological stress. Genetic manipulation is the optimal approach for assessing lncRNA function in an endogenous in vivo context, although these studies face their own challenges. There are several commonly used gene-targeting strategies that vary in their efficiency and disruption of genomic contexts (i.e., enhancers), including deletions (whole gene or promoter), inversions (whole gene or promoter), and insertions (pre-mature termination sequence or reporter) (Bassett et al., 2014). *βlinc1* is the first islet lncRNA that has been genetically disrupted at the DNA level to generate a KO mouse with no detectable *βlinc1* RNA (Arnes et al., 2016). *βlinc1* KO mice exhibited impaired glucose tolerance due to defects in insulin secretion (Arnes et al., 2016). A major caveat of gene deletion, however, is the possibility

that disruption of the genomic DNA, not the RNA, is responsible for any observed phenotype. Additional KD experiments were therefore required to show that the in vivo phenotype was not due to deletion of an enhancer for the nearby gene, Nkx2–2 (Arnes et al., 2016). As technologies for in vivo gene editing improve and become more widely implemented, it is likely that we will see a significant increase in lncRNA in vivo functional studies in complex animal models to further solidify their important functions in islet biology and diabetes.

Molecular characterization of islet lncRNAs

Tens of thousands of lncRNAs have been identified in different cell types and model organisms; however, their functions largely remain unknown. Although the tools for determining lncRNA function are technically restrictive, uncovering novel regulatory mechanisms will have the greatest impact on understanding islet function and identifying novel therapeutics for diabetes. To date, no biochemical assay has been used to directly determine the molecular mechanisms by which islet lncRNAs function, which highlights both the infancy of the field and the difficulty in implementing these techniques. The different lncRNA regulatory subtypes represent the mechanisms by which lncRNAs can act on DNA in cis or in trans, bind to complementary mRNA molecules to influence their translation, and recruit proteins to either enable or prevent their function (Wang and Chang, 2011). Based on lncRNA studies in other tissues, the most straightforward way to characterize the molecular activity of a lncRNA is to identify its interacting partners, using either protein-centric or RNA-centric approaches. Protein-centric methods, such as RNA immunoprecipitation (RIP) (Zhao et al., 2010) or cross-linking immunoprecipitation (CLIP) (Darnell, 2012), use antibodies to immunoprecipitate RNA binding protein complexes from cellular homogenate in vivo. LncRNAs stably associated with these

proteins, either directly (native RIP) or indirectly (CLIP), can be extracted and measured by quantitative PCR or unbiased identification by RNA-seq. If empirical evidence suggests lncRNA interaction with a specific protein, then these techniques are feasible. However, when functional data do not indicate a role for specific protein interactions, an RNA-centric approach is the ideal a priori strategy to probe lncRNA regulatory mechanisms. As the number of identified lncRNAs climbed exponentially, so did the need for RNA-centric biochemical purification methods. To address this technology gap, three techniques emerged almost concurrently: Capture Hybridization Analysis of RNA Targets (CHART) (Simon et al., 2011), Chromatin Isolation by RNA Purification (Chu et al., 2011), and RNA Antisense Purification (RAP) (Engreitz et al., 2013). Differences between the protocols largely pertain to the cross-linking method and probe design: CHART uses probes designed based on empirical evidence of RNaseH accessibility, whereas ChIRP and RAP both tile the whole RNA molecule, albeit with different sized oligos, 20-mer and 120-nt, respectively. The most powerful aspect of these protocols is that once RNA pull-down is successful, the readout can be tailored for tandem analyses of the DNA, RNA, or protein bound to a lncRNA in either a systematic or candidate-driven approach. For example, a comprehensive study used CHART to pull down lncRNAs *NEAT1* and *MALAT1*, followed by both DNA sequencing to identify genome-wide DNA binding sites and mass spectrometry to identify all interacting proteins (West et al., 2014).

Although identification of lncRNA molecular function is theoretically straightforward, the paucity of studies using either protein-centric or RNA-centric techniques on islet lncRNAs exemplifies the challenges associated with these tools. A major source of difficulty associated with these techniques is that they were all developed using ubiquitous lncRNAs with relatively

high endogenous expression levels, including *roX2* (CHART), *Xist* (RAP), and *roX2*, *TERC*, and *HOTAIR* (ChIRP) (Simon et al., 2011; Chu et al., 2011; Engreitz et al., 2013). Furthermore, these lncRNAs had previously characterized regulatory mechanisms, which meant positive controls were already available to trouble-shoot RNA pull-downs. Conversely, the application of these techniques to characterize the molecular mechanism of novel cell-specific lncRNAs faces challenges associated with insufficient tissue (starting material), low abundant transcripts, and unknown binding partners. LncRNA overexpression or in vitro transcription may be an option when endogenous expression is too low; however, these approaches can also introduce experimental artifacts. Due to the infancy of the lncRNA field, most of the biochemical and genetic tools used to interrogate lncRNA function have only recently been developed or are adapted from techniques used to study protein-coding genes and we are only beginning to appreciate the limits and challenges of borrowing strategies from the protein-coding world. Given the growing appreciation for lncRNAs in biology, it is likely that increased efforts will be made to adapt and optimize these technologies to enable mechanistic characterization of all functional lncRNAs, regardless of their abundance.

The discovery of lncRNAs as a novel class of tissue specific regulatory molecules has spawned an exciting new field of biology that will significantly impact our understanding of pancreas physiology and pathophysiology. As the field continues to grow, there is growing appreciation that lncRNAs will provide many of the missing components to existing molecular pathways that regulate islet biology and contribute to diabetes when they become dysfunctional. However, to date most of the experimental emphasis on lncRNAs has focused on large-scale discovery using genome wide approaches, and there remains a paucity of functional analysis. With improved

RNA-centric imaging and molecular technologies, combined with the advent of novel gene editing tools, it is likely that our knowledge of lncRNA functions in the islet will expand exponentially to rival what is currently known about canonical transcriptional regulatory programs. These advances will pave the way to a greater understanding of islet biology and enable the development novel therapies for the treatment of diabetes.

Summary and Thesis Aims

Several lines of evidence prompted us to explore the role of lncRNAs in the context of islet biology. First, genome wide association studies found that most T2DM and T1DM susceptibility loci fall outside of coding regions, which suggests a role for noncoding elements in the development of diabetes. Second, although decades of research elucidated many transcription factor networks required for pancreas development and function, the molecular mechanism mediating cell specific regulatory functions of those transcription factors remains poorly understood. Additionally, lncRNAs generally exhibit highly tissue specific expression patterns, especially compared to transcription factors expressed in multiple islet cell types or in several stages of pancreas development. Lastly, several groups have established lncRNAs as essential transcriptional regulators in development and disease (Batista and Chang, 2013). These findings formed the foundation of my thesis work, which aimed to explore novel lncRNA regulatory mechanisms that could further our understanding of islet function and dysfunction.

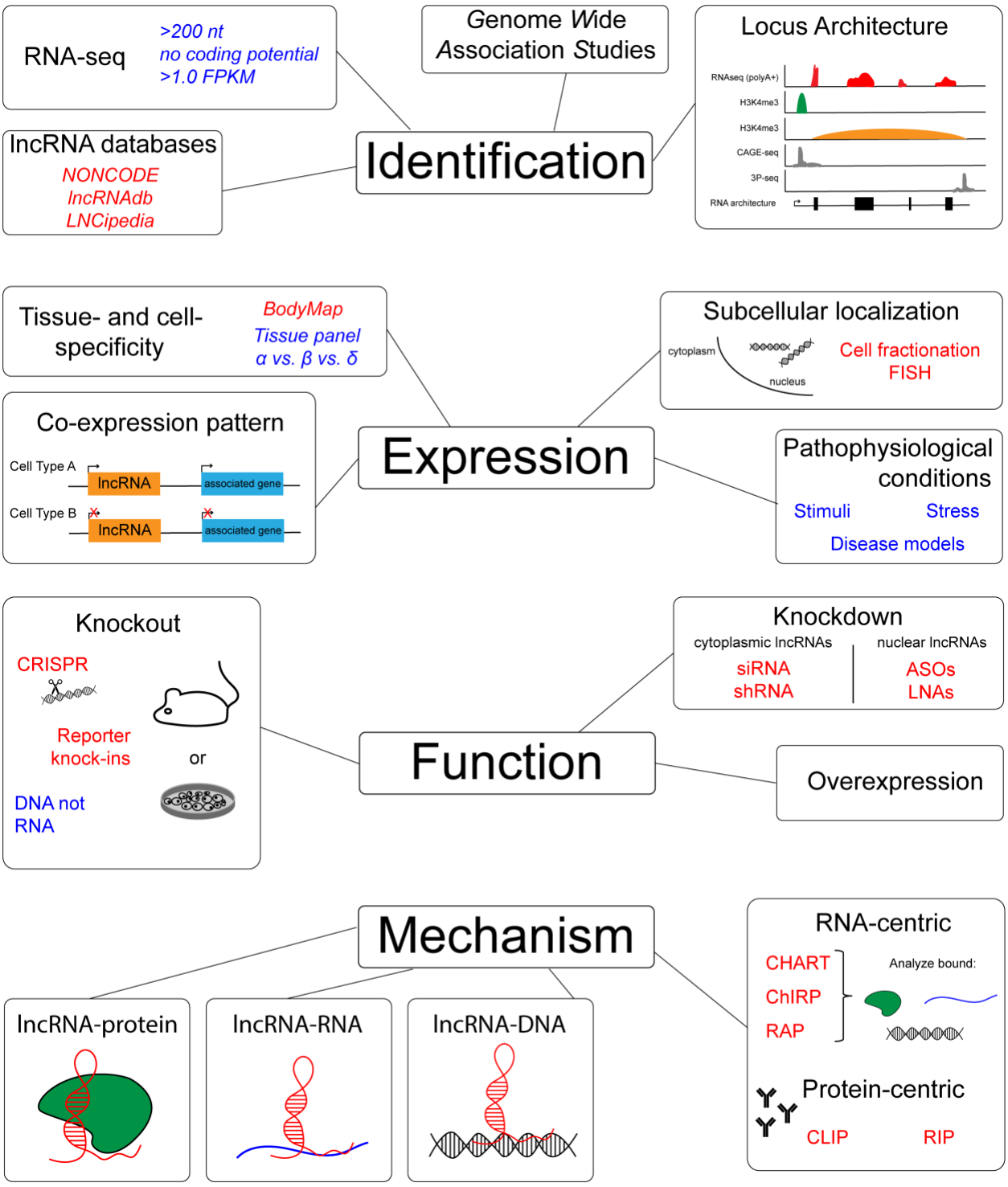


Figure 1-11. Overview of the lncRNA discovery and characterization pipeline. The lncRNA pipeline outlined in this section flows through four main phases: 1) identification of lncRNAs, 2) expression analyses, 3) functional interrogation, and 4) determination of a regulatory mechanism.

Highlighted are the tools and techniques (red), experimental parameters to consider (blue), and general strategies to characterize functional lncRNAs. 3P-seq, poly(A)-position profiling by sequencing; CAGE-seq; cap analysis gene expression; FPKM, fragments per kilobase million. Adapted from Singer and Sussel, 2018.

CHAPTER 2:

The lncRNA *Paupar* confers cell specific regulatory functions on Pax6 via alternative splicing

All experiments described in this chapter were conceived and designed by Ruth Singer with guidance from Lori Sussel. All experiments described in this chapter were performed by Ruth Singer with the following exceptions: 1) Luis Arnes performed the initial lncRNA screen in e15.5 pancreas and adult islets 2) Jiguang Wang performed the computation analyses that identified pancreatic lncRNAs 3) Yi Cui performed all RNA FISH experiments and 4) Yuqian Gao, Kristin E. Burnum-Johnson, and Charles Ansong conducted the Mass Spectrometry following CHART and analyzed the peptide data. The data presented in this chapter will be published in a manuscript of which Ruth Singer is the first author (Singer et al., in submission). Lori Sussel assisted in writing the manuscript and offered guidance, support, and mentorship throughout the project.

Summary

A current challenge in the field of vertebrate gene regulation is understanding the mechanism by which broadly expressed transcription factors exert cell specific regulatory activities. Recently, long noncoding RNAs (lncRNAs) have been identified as essential gene regulators with highly restricted gene expression. In this study, comparative transcriptome analyses between embryonic mouse pancreas and adult mouse islets identified several pancreatic lncRNAs that lie in close proximity to essential pancreatic transcription factors, including the *Pax6*-associated lncRNA *Paupar*. We demonstrate *Paupar* is enriched in glucagon-producing alpha cells where it

promotes the alternative splicing of *Pax6* to an isoform required for activation of essential alpha cell genes. Consistently, deletion of *Paupar* in mice resulted in dysregulation of PAX6 alpha cell target genes and corresponding alpha cell dysfunction. These findings illustrate a distinct mechanism by which *cis*-acting lncRNAs can contribute to cell-specific regulation of broadly expressed transcription factors to coordinate critical functions within a cell.

Introduction

Type 1 and type 2 diabetes mellitus (T1DM and T2DM) are chronic conditions with genetic, immunological, and environmental etiologies that occur due to the failure of the pancreatic islets to maintain glycemic control. Blood glucose homeostasis results from the coordinated, but opposing action of two pancreatic islet-derived hormones, insulin and glucagon. Nutrient ingestion stimulates insulin secretion from islet beta cells, which promotes glucose uptake and suppresses liver glucose production. Hypoglycemia stimulates glucagon secretion from islet cells, which promotes glucose production and release from the liver. While the majority of diabetes treatment options have focused on enhancing beta cell function to meet insulin demands, more recent work has shown that alpha cell dysfunction and hyperglucagonemia also contribute to disease pathophysiology (Unger and Cherrington, 2012; Brissova et al., 2018). Thus, a better understanding of the regulatory mechanisms required for the development and function of these highly-specialized islet endocrine cells will provide important novel information that could be applied to therapeutic treatments.

The adult endocrine cell populations are all derived from Ngn3-expressing endocrine progenitor cells, and include insulin-producing beta cells, glucagon-producing alpha cells, somatostatin-

producing delta cells, and pancreatic polypeptide-producing PP cells. A large number of transcriptional regulators that are essential for the islet cell lineage decisions have been identified and characterized. Many of these factors, including Pdx1 (Offield et al., 1996; Gao et al., 2014), Nkx2-2 (Sussel et al., 1998; Gutiérrez et al., 2017; Churchill et al., 2017), NeuroD1 (Anderson et al., 2009; Naya et al., 1997), and Glis3 (Kang et al., 2016), are expressed in several progenitor and/or islet cell populations and are continuously required in the adult for the maintenance of mature islet cell identity and function (Pan and Wright, 2011; Talchai et al., 2012; Gutiérrez et al., 2017; Ediger et al., 2017). One of the more well characterized broadly expressed islet transcriptional regulators is the paired and homeodomain transcription factor, PAX6 (St-Onge et al., 1997; Sander et al., 1997). Mice deleted for *Pax6* in pancreatic progenitor cells are born with severely reduced numbers of alpha, beta, and delta cells, suggesting that PAX6 is necessary for the development of multiple islet endocrine cell types (Ashery-Padan et al., 2004; Hart et al., 2013). Conditional deletion of *Pax6* in mature alpha or beta cells demonstrated that PAX6 is also required to maintain the identity and function of mature islet endocrine cells (Ahmad et al., 2015; Gosmain et al., 2012; Mitchell et al., 2017; Swisa et al., 2017). Furthermore, a recent study showed that within mouse beta cells, PAX6 directly activates critical beta cell genes and represses genes that specify the alternative islet endocrine cell lineages (Swisa et al., 2017). PAX6 also has essential functions in the eye (Shaham et al., 2012) and central nervous system (Manuel et al., 2015). Studies in corneal and epithelial cell lines have shown that splice variants of PAX6 bind distinct DNA motifs (Epstein et al., 1994; Chauhan et al., 2004), yet the molecular mechanism mediating the tissue-, cell-, and gene-specific regulatory functions of PAX6, and other pancreatic transcription factors, remain poorly understood.

While the majority of transcriptional regulatory proteins are expressed in many different cell and tissue types, a recently discovered class of non-protein coding RNAs (ncRNAs), called long noncoding RNAs (lncRNAs), tend to be highly tissue and cell type specific and temporally regulated (Carninci et al., 2005; Derrien et al., 2012; Guttman et al., 2011), giving them the potential to confer cell type specificity on broadly expressed transcription factors. In the pancreas, thousands of lncRNAs have been identified in human islets (Moran et al., 2012; Akerman et al., 2017; Fadista et al., 2014; Li et al., 2014), purified human alpha and beta cells (Nica et al., 2013; Bramswig et al., 2013), mouse islets (Ku et al., 2012; Motterle et al., 2017), and purified mouse alpha and beta cells (Ku et al., 2012; Benner et al., 2014), where they exhibit properties consistent with functional genes (Singer and Sussel, 2018). Studies on individual islet lncRNAs have provided evidence that lncRNAs play a role in islet function, primarily through the regulation of essential islet transcription factors, such as *HI-LNC25* on GLIS3 (Moran et al., 2012), *βlinc1* on Nkx2-2 (Arnes et al., 2016), and *PLUTO* on PDX1 (Akerman et al., 2017). These findings highlight a role for lncRNAs in islet biology and their highly restricted expression patterns further suggest that they can influence islet cell specific functions.

To identify developmentally regulated pancreatic lncRNAs, we conducted comparative transcriptome analyses of embryonic mouse pancreas and adult mouse islets and identified 572 dynamically expressed lncRNAs. Our analyses uncovered several pancreatic lncRNAs that lie in close proximity to essential pancreatic transcription factors, including the lncRNA Pax6 Upstream Antisense RNA (*Paupar*), which mapped near the *Pax6* genomic locus in mice and humans (Vance et al., 2014). Interestingly, we found that *Paupar* is enriched in glucagon-producing alpha cells where it promotes the alternative splicing of *Pax6* to an isoform

responsible for PAX6-dependent activation of essential alpha cell genes. Consistent with a role for *Paupar* in conferring the alpha cell specific functions of PAX6, deletion of *Paupar* in mice resulted in dysregulation of Pax6 alpha cell target genes and alpha cell dysfunction. Our findings illustrate how lncRNAs can modulate transcription factor activities to achieve cell-specific regulation.

Materials and Methods

Cell line and cell culture

AlphaTC1 clone 6 cells (ATCC, CRL-2934) were cultured in DMEM supplemented with 10% FBS, 15 mM Hepes, 0.1 mM non-essential amino acids, 0.02% BSA, 1.5 g/L sodium bicarbonate, 2.0 g/L glucose, and 1% antibiotic-antimycotic. The cells were passaged and maintained following standard techniques in 5% CO2 and 95% air.

Generation of *Paupar* -/- mice

The *Paupar* knockout allele was generated using BAC recombineering. First, Gibson assembly was used to clone H2BGFP (Kanda et al., 1998) and a polyA sequence and insert them together into the pL451 plasmid containing flox Neo flox (FNF) (Nam and Benezra, 2009). Short (80 bp) arms homologous to the regions flanking *Paupar* were then added to H2BGFPpA-FNF by PCR. The BAC clone (RP23-465J7; BAC-PAC Resources) was modified by insertion of the H2BGFPpA-FNF into the *Paupar* locus using Cre from SW106 cells, which was then retrieved into pMCS-DTA (a gift from Dr. Kosuke Yusa) with a 2 kb 5' arm and a 5 kb 3' rm. Positive clones were validated by PCR and DNA sequencing, and a correctly modified BAC was electroporated into mouse embryonic stem cells (129SV background) at Columbia University (Herbert Irving Comprehensive Cancer Center Transgenic Shared Resource). Potentially recombined clones were screened by PCR and two positive clones were used to generate chimeric mice that resulted in germline transmission. Chimeras were bred with FLPe transgenic mice (Jackson Laboratories) to excise the neomycin resistance cassette. *Paupar* +/- mice were backcrossed 10 generations into the C57BL/6J (Jackson Laboratories) background.

Mouse husbandry

All mice were maintained on a mixed C57BL6/129SV genetic background. Mice were group-housed in a 12-hour light/dark cycle (light between 07:00 and 19:00) with free access to water and food and maintained according to protocols (AAAQ3403) approved by the Columbia University Institutional Animal Care and Use Committee. Euthanasia was performed by CO2 inhalation. Mouse age and sex are indicated in the figure legends or methods. Genotyping for the *Paupar* WT or KO allele was performed using primers P1, P2, and P3 (Figure 2-8A). Genotyping primers are listed in Table 2-4.

LncRNA Identification

To identify novel lncRNAs, we implemented an established pipeline for identifying lncRNAs from RNA-Seq data (Arnes et al., 2018; Pefanis et al., 2015). Briefly, rRNA-depleted total RNA from e15.5 embryonic mouse pancreata and 12-week-old mouse islets was prepared using the Ribo-Zero rRNA removal kit (Epicentre). Biological replicates indicate that each RNA-seq data set was generated from individual mice (islets) or 3-4 pooled e15.5 embryonic pancreata. Libraries were prepared with Illumina TruSeq RNA sample preparation kit and then sequenced with 60 million, 2 x 100 paired reads on an Illumina HiSeq 2000 V3 instrument at the Columbia Genome Center. To reconstruct the transcriptomes, we first mapped all reads of total RNAs to the mouse reference genome (mm9) with TopHat v1.3.2 (Trapnell et al., 2012). Cufflinks v1.2.1 (Trapnell et al., 2012) was subsequently applied to assemble the whole transcriptome and to identify all possible transcripts. Then, the six transcriptomes from all samples was merged into a single gene transfer format (GTF) file with quantified gene RPKM values for each biological replicate. LncRNAs were then identified from the GTF file by removing transcripts if any of the

following criteria were met: (1) they were overlapped with genes annotated in Ensembl and not annotated as 'lincRNA', 'non_coding', 'anti-sense', '3prime_overlapping_ncrna', 'processed_transcript', 'miRNA', 'misc_RNA', or 'pseudogene'; (2) overlapped with RefSeq genes (mm9) annotated as protein coding, where the RefSeq ID began with 'NM'; (3) overlapped with pseudogenes from Pseudogene.org; (4) less than 200 nt in length; (5) predicted to have protein coding potential by the Coding Potential Assessment Tool (CPAT; coding probability >0.364); (6) gene-level maximum reads per kilobase of transcript per million mapped reads (FPKM) was less than 0.5. The 572 remaining transcripts are high-confidence pancreatic lncRNAs.

RNA extraction and quantitative RT-PCR analysis

Total RNA was extracted from tissue or cells using the RNeasy Plus Micro Kit (Qiagen) or the RNeasy Plus Mini Kit (Qiagen), depending on the sample. Purified RNA was quantified by Nanodrop (Thermo Fisher), which also determined RNA quality. 250 ng of total RNA was used to synthesize cDNA with the SuperScript III First-Strand Synthesis System (Invitrogen) and random hexamer primers. Resultant cDNA was diluted in water and 25 ng was used in each qRT-PCR reaction. Reactions were run on a Bio-Rad CFX96 Real Time System using either gene specific primers with iQ Sybr Green Supermix (Biorad) or TaqMan probes (Applied Biosystems). Expression levels were normalized to TATA-binding protein (*TBP*) and quantified using the ΔΔCT method. Sybr Green primers and AODs used are listed in Tables 2-4 and 2-5.

Cytoplasmic and nuclear RNA fractionation

AlphaTC cells were grown to confluency, detached by trypsinization, and pelleted. Half of the

pellet was used for total RNA isolation, and the other half was used for nuclear and cytoplasmic isolation using the PARIS kit (Ambion) following the manufacturer's instructions. Assessment of *Paupar* expression in each compartment was done using qRT-PCR with *Gapdh* and *Neat1* used for cytoplasmic and nuclear controls, respectively.

Single-molecule FISH (smFISH)

Oligonucleotide FISH probes were designed to target the full length of *Paupar* transcript. All probes were designed as 20-nt long, with CG content of 40-60%, no self-repeats and inner loop structures. Probes are labeled with Alexa647 at the 3' end and purchased from Integrated DNA Technologies. The sequence information of all FISH probes is provided in Table 2-4. The hybridization experiments were performed as previously described protocols (Raj et al., 2008; Lubeck and Cai, 2012; Cui et al., 2018). AlphaTCs were seeded onto collagen-coated, glass-bottom petri dishes (#1.5 thickness, MatTek) and grown to 70% confluence. Then cells were fixed with fresh 4% paraformaldehyde and quenched with 0.1% sodium borohydride. Fixed cells were permeabilized and stored with 70% ethanol until final use. FISH probes were diluted to 10 nM in hybridization solution (10% dextran sulfate, 10% formamide, 2× SSC, 0.02% RNase-free BSA, 2 mM ribonucleoside vanadyl complex). Hybridization was performed at 37°C overnight in a humid chamber. For cells co-stained with immunofluorescence, primary antibody was 1:1000 diluted and added to the hybridization solution. The next day, cells were thoroughly rinsed with 10% formamide in 2× SSC, followed by staining with secondary antibody and DAPI. The antibody information is listed in Table 2-4. Images were taken from an Olympus IX71-based single-molecule microscope equipped with 405 nm, 488 nm, 542 nm, 594 nm and 640 nm solid lasers. A 100× oil immersion objective lens (NA 1.4) was used and images were captured with

an EMCCD camera (Andor iXon Ultra 897). A z-stack scanning was performed to cover the whole cell volume (*e.g.,* 200 nm scanning step, 30-40 frames). In processing, RNA transcripts were identified at each scanning plane with Gaussian musk fitting algorithm and projected to the final reconstructed images. The processed images were subjected to counting with home-built MATLAB programs. Cell boundaries were determined and segmented based on the auto fluorescence background, in conjunction with DAPI staining. The source code for smFISH analysis is available upon request.

Immunohistochemistry

Tissues were fixed in 4% PFA in PBS overnight, washed in cold PBS, incubated in 30% sucrose, and cryopreserved. Immunofluorescence was performed on 7 mM sections. See Table 2-5 for a list of primary and secondary antibodies used. DAPI (1:1000) was applied for 15 min following a 2-hour secondary antibody incubation. Images were acquired on a Zeiss Confocal LSM 710 microscope.

Morphometric analysis

For all morphometric analysis, the entire pancreas was sectioned, and at least six evenly distributed sections were analyzed. To determine percentage hormone positive area over islet area, sections were stained with insulin, glucagon, and DAPI and five islets per section were imaged at 20X on a Zeiss Confocal LSM 710 microscope. Hormone positive area was measured using FIJI with a standardized signal threshold. Islet area was measured manually using FIJI based on morphology and guided by staining. Islets were arranged by size, and the total islet number was used to calculate the percentage represented in each size group. To determine the

percentage of proliferating alpha cells, sections were stained with glucagon, Ki67, and DAPI, and five islets per section were imaged for quantification. Data presented is from n=3 mice per genotype unless otherwise indicated.

Islet isolation

Mouse pancreatic islets were isolated by perfusion of the pancreas with Collagenase P (Roche) through the common hepatic bile duct at a concentration of 1 mg/ml in M199 medium (Invitrogen). Then pancreas was dissected out and dissociated at 37°C for 16 minutes. After several washes in M199 supplemented with 10% FBS (Gemini Bio Products), islets were then filtered (500 μm; Fisher Cat#NC0822591), hand-picked to avoid exocrine contamination, and processed for downstream applications.

Assessment of islet function

Glucose tolerance test

Mice were fasted overnight, followed by an IP injection of 2 milligram D-glucose per gram mouse weight. Tail vein blood samples were collected at 0, 15, 30, 60, 90, and 120 minutes after the injection and glucose concentration was determined using the Accu-Chek Compact Plus Blood Glucose Meter (Roche).

Insulin tolerance test

Mice were fasted for 5 hours starting at 9 am. Insulin (Humalog in PBS) was injected IP at 0.75 units per kg mouse weight. Tail vein blood samples were collected at 0, 15, 30, 60, 90, and 120 minutes after the injection and glucose concentration was determined using the Accu-Chek

Compact Plus Blood Glucose Meter (Roche). To measure plasma glucagon during an insulin tolerance test, blood was collected into heparinized tubes immediately before, and 30 minutes after, an insulin injection. Plasma was separated from whole blood by centrifugation and glucagon concentration was measured by ELISA (Mercodia).

Glucose stimulated glucagon secretion assays

After isolation, duplicate samples of 20 islets per mouse were cultured overnight at 37°C in resting medium (RPMI 1640 with 10% FBS, 1% P/S, and 5.6 mM D-glucose). The next day, islets were transferred into modified Krebs buffer with 0.1% BSA and 20 mM D-glucose for 30 minutes to equilibrate. Islets were then transferred to Krebs buffer with 20 mM glucose for 30 minutes, followed by 2 mM glucose for 30 minutes. Lastly, islets were transferred to either Krebs buffer with 20 mM glucose or 10 mM L-Arginine for 30 minutes. Supernatant was collected following each treatment. Islets were then sonicated in 50 µl of lysis buffer (TE with 0.1% SDS) and lysate was used to determine DNA concentration and measure glucagon content. Glucagon concentration in supernatant and islet lysate was measured by glucagon ELISA (Mercodia).

RNA sequencing (RNA-seq)

Total RNA from *Paupar* WT and *Paupar* KO mice was converted into cDNA libraries (TruSeq RNA sample preparation kit, Illumina) according to manufacturer's instructions. Sequencing was performed to a depth of 30 million, single-end 100 nt reads in three biological replicates per genotype. Reads were aligned to the mouse genome (mm9) using HISAT2 v2.1.0 (Kim et al., 2013). Aligned reads were assigned to genes using annotations from Ensembl

(Mus_musculus.GRCm38.73.gtf) and HTseq-count v0.10.0 (Anders et al., 2015). Differential expression across cohorts was assessed using DESeq2 v1.18 (Love, Huber, and Anders, 2014). All samples had RNA integrity (RIN) values > 8.0 as determined with Agilent Bioanalyzer 2100. Complete RNA-seq data are available through GEO accession numbers: GSE122033 (embryonic pancreas and islet lncRNAs) and GSE121884 (Paupar WT and KO islets).

RNA interference

The day before transfection, 500,000 alphaTCs were plated in each well of a 12-well plate. Cells were transfected in triplicate with 40 nM *Paupar* or control ASOs (IDT) using Lipofectamine 2000, per the manufacturer's instructions (Thermo Fisher). Cells were harvested for RNA 48 hours post transfection. ASO sequences are listed in Table 2-4.

Capture Hybridization Analysis of RNA Targets (CHART) enrichment and analysis

CHART enrichment experiments were performed as previously described (Simon et al., 2011). Briefly, CHART extract was prepared from approximately 5 x 10^7 alphaTC cells per CHART reaction and hybridized with 120 pmol biotinylated capture oligonucleotides (COs) (Table 2-4) overnight with rotation at room temperature. Complexes were captured using MyOne Streptavidin C1 beads (Invitrogen). Bound material was stringently washed and eluted using RNase H (New England Biolabs) for 15 minutes at room temperature. For RNA analysis, CHART-enriched material was incubated with XLR buffer (final concentrations: 2 mg/mL Proteinase K (Ambion), 33.3 mM Tris pH 7.2, 0.33% SDS, 16.7 mM EDTA) at 55°C for 1 hour and 65°C for 1 hour to reverse cross-links. RNA was purified using Trizol LS according to the manufacturer's instructions. qPCR analysis utilized iTaq Universal SYBR Green One-Step Kit to

quantify *Paupar* enrichment compared to control genes. For analysis of CHART enriched proteins, eluate was precipitated using trichloroacetic acid and resuspended in SDS (4.25%), Tris pH 8.8 (529 mM), EDTA (64 mM) and β-mercaptoethanol (1.37 M). Samples were incubated at 98°C for 30 minutes and 65°C for 2 hours to reverse cross-links. Samples (including 1%, 2.5%, and 5% input) were resolved by SDS-PAGE and transferred to nitrocellulose membranes. Membranes were incubated with antisera to detect PAX6 (Biolegend) and bands visualized with SuperSignal West Pico PLUS Chemiluminescent Substrate.

CHART-MS

Following CHART pulldown, samples were precipitated overnight at -20 ℃ with 4 volumes of cold acetone. The precipitated sample was centrifuged for 20 min at 16,000 x g at 4 ℃ and the supernatant discarded. The remaining pellet was washed with 1 mL cold acetone, centrifuged for 5 min at 16,000 x g at 4 ℃, and the supernatant discarded. The remaining pellet was dried under vacuum. The protein pellet was digested to peptides using the TFE digest protocol (Wang et al., 2005) and peptide samples were analyzed by mass spectrometry. For LC-MS/MS analysis, 5 µl of 0.1 µg/µl of peptides per sample were analyzed by reverse phase separation (C18) using a Waters nanoEquityTM UPLC system interfaced with a QExactive Plus Orbitrap mass spectrometer (Fort et al., 2018). LC-MS/MS datasets were converted to peak lists (DTA files) using the DeconMSn software (Mayampurath et al., 2008) and searched with MS-GF+ software (Kim, Gupta, and Pevzner, 2008; Kim and Pevzner, 2014) against Uniprot/SwissProt *mus musculus* database. The identified spectra were filtered based on their MSGF+ QValue score, which represents the false discovery rate of peptide-spectrum-matches (PSMs). PSMs with QValue less than 0.01 were retained. Proteins were quantified using the spectral counting

quantification approach (Wang et al., 2008). Downstream analyses identified *Paupar* binding proteins that were enriched > 2-fold over controls, excluding proteins found in > 75% of control experiments listed in the Contaminant Repository for Affinity Purification (CRAPome; Mellacheruvu et al., 2013). We used Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) (Szklarczyk et al., 2017) to identify known and predicted protein-protein interaction networks among our *Paupar*-interacting proteins.

ChIP-qPCR

The ChIP-IT High Sensitivity kit (Active Motif) was used according to the kit instructions. Briefly, 48 hours post ASO transfection (after ensuring efficient *Paupar* KD), cells were crosslinked, lysed, and sonicated to obtain DNA fragments with an average length of 200–1200 bp. Supernatant containing DNA-protein complexes was used for immunoprecipitations reactions consisting of 30 µg chromatin and 4 µg rabbit anti-PAX6 antibody (Biolegend) or rabbit IgG control antibody (included in kit). Immunoprecipitated chromatin was collected using protein G agarose beads, then washed, eluted, and purified according to kit instructions. ChIP-qPCR were performed using the iQ Sybr Green Supermix (Biorad) and primers against *MafB* and *Glucagon* promoter regions. Data were normalized to the input signal and IgG values.

Quantification and Statistical Analyses

Graphs were generated in GraphPad Prism 7 and statistical analyses were performed using Graphpad Prism 7. Statistical parameters including the value of n, statistical test used, and significance (p-value) are reported in the figures and their legends. For studies involving mouse tissues, replicates refer to samples derived from different mice, unless otherwise specified. For

studies involving cell culture, replicates refer to transfections performed on different days. Unpaired two-tailed t tests were used to assess significance when comparing qRT-PCR expression values of 2-3 genes between two conditions. Linear regression analysis was used to determine if two variables were significantly correlated to each other. The Fisher's exact test was used to examine the significance of the association (contingency) between two types of classifications. For differential expression of global measurements (RNA-seq), the DESeq2 software (Love et al., 2014) generated adjusted p-values using the Benjamini-Hochberg procedure to correct for multiple-hypothesis testing. The accession numbers for the sequencing data reported in this chapter are GSE122033 (embryonic pancreas and islet lncRNAs) and GSE121884 (*Paupar* WT and KO islets).

Results

Identification of developmentally regulated lncRNAs in the mouse pancreas

To identify temporally and spatially regulated lncRNAs in the pancreas, we performed RNA-sequencing (RNA-seq) on adult mouse islets and embryonic day 15.5 (e15.5) mice pancreas (EP) (n=3 per stage) to obtain 60 million paired-ended reads per sample. We analyzed these datasets using a computational pipeline (Pefanis et al., 2015) designed to identify putatively functional mouse pancreatic lncRNAs (Figure 2-1A) (Singer and Sussel, 2018). A set of stringent criteria was used to define 2728 pancreatic lncRNAs: (1) > 200 nucleotides (nt) in length; (2) no overlap with protein coding regions; (3) no overlap with pseudogenes (Karro et al., 2006); and (4) low predicted coding probability (Wang et al., 2013) (Figure 2-1B). We then filtered out genes with an FPKM < 0.5 to enrich for lncRNAs amenable to molecular analyses (Figure 2-1C). These parameters yielded 572 high-confidence pancreatic lncRNAs that clustered according to developmental stage (Figure 2-1D). Remarkably, comparative transcriptomics between the two stages showed that approximately half (279 or 48.7%) of all pancreatic lncRNAs were developmentally regulated: 108 lncRNAs were significantly enriched in EPs and 171 lncRNAs were significantly enriched in adult islets (Figure 2-1E).

Previous studies have shown that subsets of lncRNAs regulate, or are co-regulated with, nearby protein-coding genes (Cabili et al., 2011; Moran et al., 2012). To gain insight into the function of these lncRNAs, we utilized Genomic Regions Enrichment of Annotations Tool (GREAT) (McLean et al., 2010) to identify nearby protein-coding genes (Figure 2-1F). Several lncRNAs mapped within 5 kb of a nearby transcriptional start site (TSS), suggesting they might function as bidirectional lncRNAs (Mercer, Dinger, and Mattick, 2009), while the majority of lncRNAs were

located in intergenic regions (> 50 kb from a TSS) (Figure 2-2A). Gene ontology analyses of all neighboring genes showed a significant enrichment of genes involved in "endocrine pancreas development" and "pancreas development" (Figure 2-1G), including nine essential pancreas transcription factors: *Pax6*, *Foxa2*, *Nkx6.1*, *NeuroD1*, *Hes1*, *Foxo1*, *Mnx1*, *Meis2*, and *Hnf6* (Table 2-1).

Paupar is a nuclear lncRNA enriched in pancreatic alpha cells

One of the more differentially regulated lncRNAs (~40x higher in islets compared to EP; Figure 2-1E) was the *Pax6* Upstream Antisense RNA (*Paupar*), a previously characterized lncRNA located near the *Pax6* locus (Vance et al. 2014) with an orthologous transcript in human islets (HI-LNC101) (Moran et al., 2012). *Paupar* is a 3482 bp gene transcribed 8 kb upstream and antisense from Pax6 and contained within the first intron of *Pax6os1*, an antisense RNA with no known function (Figure 2-3A). *Paupar* lies in a syntenically conserved region on chromosome 2 in mice and chromosome 11 in humans (LiftOver). *Paupar* is also highly conserved at the nucleotide level across mammals (89% in humans; Table 2-1), both in its gene body and regions located at the putative promoter region (Figure 2-3A), similar to many functional lncRNAs (Carninci et al. 2005). Comparative sequence analysis of *Paupar* across species did not reveal any conserved ORFs, reinforcing the computational prediction that *Paupar* lacks protein-coding potential (PhyloCSF -6.5127; CPC 0.2666; CPAT 0.2086).

Paupar was previously identified as a single exonic lncRNA in the neuroblastoma N2A cell line, where it appears to have PAX6-dependent and independent functions (Vance et al., 2014). To more globally characterize *Paupar* expression, we performed extensive analysis of published

RNA-seq datasets from 25 different mouse tissues (Lin et al., 2014) and found that *Paupar* was expressed exclusively in the pancreas, eye, and brain (Figure 2-4A). These datasets also confirmed the findings of our initial screen, showing enrichment of *Paupar* in adult islets compared to embryonic pancreas (Figure 2-1E). Furthermore, expression data from alphaTC and MIN6 cell lines suggested restricted expression in islet alpha cells (Figure 2-4A). Surprisingly, in contrast to what was reported in N2A neuroblastoma cells, in the context of islet alpha cells, *Paupar* appears to contain three exons and two introns (Figures 2-4A-C).

We confirmed the RNA-Seq data using qRT-PCR on RNA from 11 different mouse tissues to demonstrate that *Paupar* was significantly enriched in adult mouse islets, compared to the eye (~3.5-fold) and brain (~25-fold) (Figure 2-3B). Expression analysis in whole pancreata at several embryonic and postnatal developmental time points also revealed that onset of *Paupar* expression occurs between postnatal day 7 (P7) and P14 (Figure 2-3C), which is a critical window of postnatal development associated with the functional maturation of endocrine cells (Nishimura et al., 2006; Hang and Stein, 2011). Given that PAX6 has known regulatory roles in both alpha and beta cells (Gosmain et al., 2010; Swisa et al., 2017), we wanted to confirm the alphaTC and MIN6 expression data that suggested *Paupar* expression was restricted to the alpha cell lineage. Analysis of published RNA-seq datasets from FACS purified mouse alpha and beta cells (DiGruccio et al., 2016) confirmed *Paupar* was enriched 3-fold in alpha cells compared to beta cells (Figure 2-3D), in contrast to *Pax6,* which is expressed equally in both cell types (Figure 2-3E). Expression levels of *MafB* and *MafA,* essential alpha and beta cell transcription factors, respectively, are shown to demonstrate the purity of each dataset (Figure 2-3E).

LncRNA localization within a cell can also yield important mechanistic insight; nuclear lncRNAs often regulate transcription and pre-mRNA processing, while cytoplasmic lncRNAs more likely influence mRNA stability and translation (Batista and Chang, 2013). Expression analyses on alphaTCs following cell fractionation revealed a significant enrichment of *Paupar* in the nucleus compared to the cytoplasm (Figure 2-3F). Nuclear localization of *Paupar* in alphaTC cells was confirmed by single-molecule fluorescent in situ hybridization (smFISH) using oligonucleotide probes targeting the full length *Paupar* transcript. In comparison, glucagon protein is localized to the cytoplasm (Figure 2-3G). Analysis of *Paupar* smFISH images determined that the average copy number for *Paupar* is 54 ± 31 transcripts per alphaTC cell (Figure 2-4D). Taken together, these results demonstrate *Paupar* is a conserved lncRNA predominantly localized to the nuclei of mature pancreatic alpha cells.

Paupar regulates Pax6 alpha cell target genes

The expression profile of *Paupar* prompted us to investigate its regulatory function in alpha cells. Three unique sets of antisense oligonucleotides (ASO) successfully downregulated *Paupar* RNA in alphaTC cells by an average of 57% (Figure 2-5A). In contrast to what has been previously reported in N2A cells, we observed no change in *Pax6* expression following *Paupar* knockdown (Figure 2-5A), suggesting that *Paupar* RNA does not influence *Pax6* transcription. We were surprised then, to discover that *Paupar* knockdown led to the downregulation of several canonical *Pax6* alpha cell target genes, including *Gcg, MafB, Arx*, and *NeuroD1* (Gosmain et al., 2010) (Figure 2-5A). Furthermore, the amount of *Paupar* KD was positively correlated with the reduction in both *Gcg* (Figure 2-5B) and *MafB* (Figure 2-5C) expression. These findings suggest that *Pax6*-mediated regulation of alpha cell genes is sensitive to the level of *Paupar* present in

alphaTCs. *Foxa2*, an alpha cell gene that is not a target of PAX6, was not regulated by loss of *Paupar*, further suggesting the activity of *Paupar* is restricted to PAX6 target genes (Figure 2-5A).

Paupar lncRNA interacts with nuclear proteins involved in alternative splicing

To investigate the molecular mechanism by which *Paupar* regulates alpha-cell Pax6 target genes, we performed Capture Hybridization Analysis of RNA Targets (CHART) (Simon et al., 2011) to identify the *Paupar* interactome. We used two unique sets of biotinylated capture oligos (COs) to pull down *Paupar*, in addition to scrambled and *Paupar* sense control COs. Both sets of *Paupar* COs retrieved *Paupar* RNA from alphaTC nuclear extract (Figure 2-5D). Importantly, *Paupar* COs did not pull down *Gapdh* or *TBP*, nor did control COs retrieve *Paupar* (Figures 2-5D and 2-5E). The lack of *Paupar* enrichment with sense COs, which cannot directly target the RNA but have sequence identity to the DNA locus, demonstrates that the CHART-based enrichment is RNA mediated (Figure 2-6A).

To identify *Paupar*-interacting proteins, we performed CHART followed by mass spectrometry (CHART-MS), and identified 56 *Paupar* binding proteins that were enriched > 2-fold over control COs (Table 2-2). Surprisingly, unlike in N2A cells, we were not able to detect an interaction between *Paupar* and PAX6 protein (Figure 2-6B). Analysis of the *Paupar* interacting proteins using a database of known and predicted protein-protein interaction networks (Search Tool for the Retrieval of Interacting Genes/Proteins (STRING); Szklarczyk et al., 2017) identified 25/56 proteins that interact with at least one other *Paupar* interacting protein (Figure 2-5F). These protein complexes were significantly enriched in several pathways (Figure 2-5G),

including RNA splicing (Figure 2-5F, blue circles), regulation of gene expression (Figure 2-5F, red circles), and DNA binding (Figure 2-5F, green circles). Since the largest complex of interacting proteins contained canonical regulators of alternative splicing, we compared the set of *Paupar* interacting proteins with those retrieved by *NEAT1* and *MALAT1,* two nuclear lncRNAs with established roles in alternative splicing (West et al., 2014). This identified 50.6% and 51.7% overlap with *NEAT1* and *MALAT1* interacting proteins, respectively (Figure 2-6C). In contrast, similar comparison to proteins or lncRNAs that are not involved in alternative splicing mechanisms, MAFA (Scoville et al., 2015), NEUROD1 (Romer et al., 2018) and *Xist* (Chu et al., 2015), showed limited overlap (Figure 2-6C).

Consistent with a direct interaction between *Paupar* and 8 of the 12 annotated serine and arginine rich splicing factors (SRSFs) (Figure 2-5F; Table 2-2), we used a computational tool (RBPmap; Paz et al., 2014) to demonstrate that the full-length *Paupar* transcript had 1685 predicted SRSF binding sites above a stringent z-score threshold (> 2). Notably, several of the top SRSF sites with the highest z-scores were located within a region of high sequence conservation across placental mammals (Figure 2-5H, grey rectangle). To address the possibility that SRSF binding occurs because *Paupar* is a spliced transcript, we compared the z-scores for *Paupar* SRSF binding sites to *Nkx2-2*, an mRNA that is spliced but does not regulate RNA splicing, as well as *Malat1*, a lncRNA that is not spliced but regulates SRSF-mediated RNA splicing. We found that *Paupar* had more SRSF binding sites with higher z-scores than both *Nkx2-2* and *Malat1* (Figure 2-5H). SRSF binding sites for the full ~7 kb *Malat1* locus is shown in Figure 2-6D. Taken together, these findings suggest that *Paupar* functions in the alpha cell to regulate alternative splicing.

Paupar promotes the alternative splicing of Pax6 to the isoform required for activation of Pax6 alpha cell target genes

Previous studies have demonstrated that the regulatory role of *Paupar* in N2A cells is partially mediated through PAX6 (Vance et al., 2014; Pavlaki et al., 2017). The identification of interactions between *Paupar* and the SRSF family of proteins, and the knowledge that *Pax6* has two well-characterized isoforms with distinct regulatory functions (Epstein et al., 1994; Chauhan et al., 2004; Kiselev et al., 2012; Sasamoto et al., 2017) (Figure 2-7A), led us to examine a possible role for *Paupar*-mediated alternative splicing of *Pax6*. The two major *Pax6* isoforms, termed *Pax6* (Figure 2-7A, red lines) and *Pax6 5a* (Figure 2-7A, blue lines), differ from each other by an alternatively spliced exon, "5a", that adds 14 amino acids to the paired DNA-binding domain of PAX6 protein and alters Pax6 DNA binding recognition (Kiselev et al., 2012). Since PAX6 regulates distinct target genes in alpha versus beta cells (Gosmain et al., 2010; Gosmain et al., 2012), it is possible that its cell-specific regulatory activities could be mediated through its different isoforms. Consistently, computational analysis of alternative splicing events in alpha versus beta cell transcriptomes (DiGruccio et al., 2016) identified the *Pax6 5a* isoform as significantly enriched in alpha cells ($p < .0001$; Table Appendix-1). While non-quantitative RT-PCR analysis detects both isoforms in alpha cells (Figure 2-7B), we could only quantify expression of *Pax6 5a*, given that amplification of the longer *Pax6* isoform produces both isoform products, which confounds qRT-PCR analysis. To directly test whether *Paupar* promotes alternative splicing of *Pax6*, we performed qRT-PCR on RNA from *Paupar*-deficient alphaTCs. Strikingly, while *Paupar* KD did not induce a change in total *Pax6* mRNA levels (Figure 2-5A), we did observe a significant and specific reduction (45%) of *Pax6 5a* (Figure 2-7C). We also determined that *Paupar*-mediated regulation of *Pax6* splicing occurs directly

though RNA-RNA interaction, rather than indirectly through SR splicing factors, since the association between Paupar and *Pax6* RNA that could be partially ablated by RNase treatment (Figure 2-7D).

A study in non-pancreatic cell lines that stably expressed *Pax6* or *Pax6 5a* showed that the two proteins have different DNA binding specificities (Kiselev et al., 2012). To determine whether *Paupar* mediated-alternative splicing of *Pax6 5a* is necessary for the PAX6 activation of alpha cell target genes, we performed PAX6 ChIP in alphaTC cells with normal or reduced amounts of *Paupar* and assessed changes in PAX6 occupancy on the promoters of *Gcg* and *MafB*, two genes directly activated by PAX6 (Gosmain et al., 2007; Gosmain et al., 2010). As expected, in control alphaTCs there was a greater than 30-fold enrichment of PAX6 over IgG on *Gcg* and *MafB* regulatory DNA (Figure 2-7E). Strikingly, *Paupar* KD induced a 34% reduction in PAX6 occupancy on the *Gcg* promoter and a 58% reduction in PAX6 occupancy on the *MafB* promoter, directly demonstrating that *Paupar* is required for normal PAX6 binding and activation of *Gcg* and *MafB* in alphaTCs (Figure 2-7E). Taken together, these results show that *Paupar* confers the alpha-cell specific activating function of *Pax6 5a* via alternative splicing.

Paupar knockout mice have impaired alpha cell development and function

To determine whether *Paupar*'s regulation of *Pax6* isoform selection affected alpha cell function in vivo, we generated *Paupar* null (KO) mice by replacing the endogenous *Paupar* locus with the histone-fusion GFP (H2B:GFP) reporter gene (Kanda et al., 1998) (Figure 2-8A). Expression analysis of islets from *Paupar* KO mice confirmed the complete loss of *Paupar* RNA (Figure 2-

8B), and immunofluorescence analyses indicated that the GFP reporter recapitulated endogenous *Paupar* expression specifically in the glucagon-producing alpha cells (Figures 2-8C and 2-8D).

Paupar KO mice are viable, fertile, and indistinguishable from their WT littermates with respect to weight (Figure 2-10A), *ad libitum* blood glucose (Figure 2-10B), and glucose tolerance (Figures 2-10C-F). These findings are not surprising given the restricted expression of *Paupar* in alpha cells; several studies have shown that mice are highly resistant to perturbations in alpha cell function (Furuta et al., 2001; Shiota et al., 2003; Heller et al., 2004; Hancock et al., 2010; Wilcox et al., 2013). We therefore utilized an in vivo insulin tolerance test (ITT) to induce hypoglycemia, a physiological trigger for glucagon secretion from alpha cells. Strikingly, at all assayed time points, 6-week-old *Paupar* KO mice were significantly hypoglycemic compared to WT mice, suggesting KO mice were less efficient at returning to baseline blood glucose levels (Figure 2-10A and 2-10B). To differentiate between enhanced insulin sensitivity and an inability to properly respond to hypoglycemia, we measured plasma glucagon before and during an ITT. Thirty minutes after an insulin injection, *Paupar* KO mice had ~3-fold decrease in plasma glucagon relative to WT mice (Figure 2-10C), demonstrating that *Paupar* KO mice secrete significantly less glucagon in response to insulin-induced hypoglycemia.

The physiological response to an insulin challenge is complex and involves several non-pancreatic tissues, including liver, muscle, and fat. Although *Paupar* is not expressed in those tissues (Figure 2-4A), to eliminate the possibility that the impaired response of *Paupar* KO mice to hypoglycemia was due to increased uptake of glucose by the peripheral tissues, we performed ex vivo glucagon secretion assays. Cultured islets from 6-week-old *Paupar* KO mice secreted ~60% less glucagon than WT mice in response to low glucose (2 mM) and ~47% less glucagon

than WT mice in response to 10 mM arginine, both potent glucagon secretagogues (Gerich, Charles, and Grodsky, 1974) (Figure 2-9D). This blunted glucagon secretion could be due to several factors, including impaired glucose sensing, membrane depolarization, glucagon production, and granule exocytosis. To distinguish between these defects, we began by measuring total islet glucagon content and found that *Paupar* KO islets had dramatically less (56%) total glucagon than controls (Figure 2-9E). These findings demonstrate that decreased glucagon content in *Paupar* KO islets contributes to impaired alpha cell physiological function.

Consistent with the physiological phenotype, morphometric analysis on 6-week-old *Paupar* WT and KO pancreata revealed that while alpha cells made up 13.74% ± 0.79 of *Paupar* WT islets (Figure 2-9F), they comprised only 6.11% ± 0.71 of *Paupar* KO islets (Figure 2-9G), corresponding to an average 2.25-fold decrease in the alpha cell population (Figures 2-9F-H). We also observed a modest increase in beta cell area relative to islet area in *Paupar* KO mice (79.4% ± 1.92) compared to WT mice (Figures 2-9I-K). There was no measurable difference in average islet size between *Paupar* WT and KO mice (Figures 2-9L-N). Examination of islet morphology in 7-month-old WT and KO mice showed that *Paupar* KO mice had significant alpha cell (Figures 2-11A and 2-11B) and islet hyperplasia (Figures 2-11C) due to increased alpha cell proliferation (Figures 2-11D-F). This compensation phenomenon has been well documented in several mouse models of alpha cell dysfunction (Gelling et al., 2003; Conarello et al., 2007; Hayashi et al., 2009; Courtney et al., 2013; Solloway et al., 2015) and provides further evidence that *Paupar* functions similarly to canonical alpha cell regulatory proteins. Taken together, these results demonstrate that *Paupar* is required for normal alpha cell development and function.

Paupar regulates essential alpha cell genes *in vivo*

To validate the in vitro studies which used an RNA knockdown approach to disrupt *Paupar* activity in alphaTC cells, we performed global transcriptome analyses on 6-week-old isolated islets from *Paupar* WT and KO mice to reveal 3106 dysregulated genes ($p < .05$) across cohorts (Figure 2-12A). Consistent with the alpha-cell specific phenotype, the differentially expressed genes (DEGs) represented > 10% of the documented alpha cell enriched genes (DiGruccio et al., 2016) (Figure 2-12B). Furthermore, the DEGs contained many factors required for alpha cell function and many known PAX6 alpha cell targets, including transcription factors (*Arx*, *MafB*, *Irx1*, and *Irx2*) (Table 2-3), voltage-gated ion channels (*Slc41a2*, *Kcnq2*, *Kcnip3)*, and exocytotic machinery (*Sytl5* and *Sytl2*) (Figure 2-12C). In addition, the majority (63/75) of the dysregulated alpha-cell specific genes were downregulated in *Paupar* KO mice, indicating that *Paupar* predominantly regulates gene activation in alpha cells. Also consistent with our in vitro studies, in vivo transcriptome analyses of *Paupar* KO islets did not show dysregulation of total *Pax6* mRNA, suggesting *Paupar*-mediated regulation of PAX6 alpha-cell target genes is downstream of *Pax6* transcription. Given the low abundance of alpha cells in mouse islets, and that the RNA-seq was performed on whole islets, the downregulation of *Pax6 5a* specifically in alpha cells could not be assessed. Of note, *Paupar* KO islets had normal levels of *Pax6os1*, as well as genes within 100 kb of *Paupar*, suggesting that the *Paupar* KO mouse phenotype is not due to deletion of an important DNA regulatory element. Based on these cumulative studies, we propose a model that *Paupar* promotes alpha cell function via alternative splicing of the *Pax6 5a* isoform required for activation of alpha cell target genes (Figure 2-13).

Discussion

During the past 20 years, a major focus of diabetes research has been directed towards identifying the complex transcription factor networks required for the specification and function of pancreatic islet cells. Yet, how these broadly expressed transcription factors acquire unique regulatory functions at different stages of pancreas development and in different islet cell types remains poorly understood. In this study we uncover a novel mechanism by which the lncRNA, *Paupar*, confers cell specific regulatory function on the essential pancreatic transcription factor, PAX6, by promoting the alternative splicing of *Pax6* to the *5a* isoform that is required for the activation of downstream alpha cell target genes (Figure 2-13). We have shown that the loss of *Paupar* blunts the production of *Pax6 5a* isoform, causing diminished activation of PAX6 target genes and impaired glucagon-mediated glucose homeostasis (Figure 2-13). These findings uncover a novel layer of islet gene regulation and provide further evidence that lncRNAs are fundamental players in islet development and function.

Paupar was first identified in neuroblastoma N2A cells as a single exon lncRNA that regulated genes independently and through direct interaction with PAX6. Surprisingly, our expression analyses showed that *Paupar* was most highly enriched in pancreatic islets compared to all other tissues examined, including the eye, brain, and N2A cells. Furthermore, unbiased mapping of the adult islet transcriptome showed evidence of splicing within the *Paupar* locus; cloning and sequencing confirmed that the *Paupar* transcript has three exons and two introns. We also discovered that *Paupar* does not directly interact with PAX6, nor does it regulate *Pax6* expression at the transcript level. These discrepancies between our findings and previous studies in N2A cells possibly reflect tissue-specific regulation and functional activities of *Paupar*.

Furthermore, although *Paupar* is also expressed in the mouse eye and brain, we did not observe any gross abnormalities or phenotypes associated with these tissues in *Paupar* KO animals. It is also not likely that the loss of *Paupar* in the brain contributes to the alpha cell phenotype, since we observed impaired glucose-induced glucagon secretion and a significant decrease in glucagon content in isolated islets of *Paupar* KO mice. Lastly, *Paupar*-mediated regulation of alpha cell genes in vitro was consistent with in vivo phenotypes.

Within the pancreas, *Pax6* is expressed in several islet cell types, while *Paupar* expression is restricted to maturing alpha cells. Intriguingly, the onset of *Paupar* pancreatic expression between P7 and P14 corresponds to a critical developmental window during which cells acquire mature transcriptional profiles. For example, during this postnatal window the transcription factors *MafA* and *MafB* become restricted to mature beta and alpha cells, respectively. These findings, along with the dramatic reduction in alpha cells seen in 6-week-old *Paupar* KO mice, and experiments showing reduced PAX6-activation of *MafB* in *Paupar*-deficient alphaTC cells, demonstrate that *Paupar* is required for the differentiation of mature alpha cells by promoting PAX6-mediated activation of *MafB.*

Previous studies have shown that PAX6 functions as a transcriptional activator and a repressor; however, the mechanism that mediates this dual capability in a single cell type is unknown. The discovery that *Paupar* expression specifically in alpha cells promotes the alternative splicing of *Pax6* to an isoform with altered DNA binding specificity raises the interesting possibility that the different PAX6 isoforms have unique functions. Since were able to detect both the *Pax6* and *Pax6 5a* isoforms within alpha cells, it is possible that the ratio of each isoform is crucial for

proper gene regulation. This idea is supported by our finding that reduction of the *Pax6 5a* isoform in *Paupar*-deficient alpha cells corresponded to reduced expression of several canonical PAX6-activiated alpha cell genes, including *glucagon* and *MafB* (Gosmain et al., 2007; Gosmain et al., 2010), while *ghrelin*, the only gene known to be repressed by PAX6 in alpha cells (Ahman et al., 2015), was not upregulated in *Paupar* KO islets or in *Paupar*-deficient alphaTC cell line (Figure 2-12 and data not shown).

The majority of *Paupar* molecular analysis was performed in the alphaTC cell line, which is an appropriate, but imperfect model for endogenous alpha cells. A current challenge in the alpha cell field is the paucity of data characterizing transcription factor binding and chromatin structure in endogenous mouse alpha cells. This is at least partially due to the relatively low proportion of alpha cells in the mouse islet; it is difficult to collect sufficient material from isolated alpha cells for biochemical analysis and data generated from whole islet preparations largely represent the exceedingly abundant beta cell population. However, with the continual development of novel technologies to perform high throughput molecular analyses on low abundant cell populations, these barriers will soon be reduced. In the future, the ability to perform ChIP-Seq on purified islet alpha and delta cells will likely highlight the distinct islet cell specific regulatory mechanisms to inform both normal islet cell type specification, and the dedifferentiation and reprogramming events that have recently been associated with diabetic islets in mice and humans (Talchai et al., 2012; Lu et al., 2018).

While several lncRNAs have been shown to generally influence alternative splicing through their direct interactions with splicing factors (reviewed in Romero-Barrrios et al, 2018), this study

demonstrates for first time a specific interaction between a cell-restricted lncRNA and its *cis*-related gene that in turn influences the production of distinct protein isoforms. In particular, the presence of *Paupar* in islet alpha cells skews the alternative splicing of *Pax6* to favor a PAX6 isoform required for alpha cell specific gene activation. These results highlight an important mechanism through which tissue restricted lncRNAs influence transcription factor target selection to confer cell specific activities. The identification and characterization of additional cell restricted lncRNAs will be instrumental in determining the extent of this gene regulatory mechanism. We predict that additional evidence of *cis* lncRNA-regulation of tissue-specific mRNA splicing will emerge, since many lncRNAs function in the same pathway as their neighboring gene, even though there is often no evidence of co-regulation at the transcriptional level. In summary, our extensive molecular and functional characterization of an alpha cell enriched lncRNA suggest that lncRNAs could represent important tissue and/or cell restricted therapeutic targets to regulate the production and function of cell-specific isoforms of more widely expressed proteins.

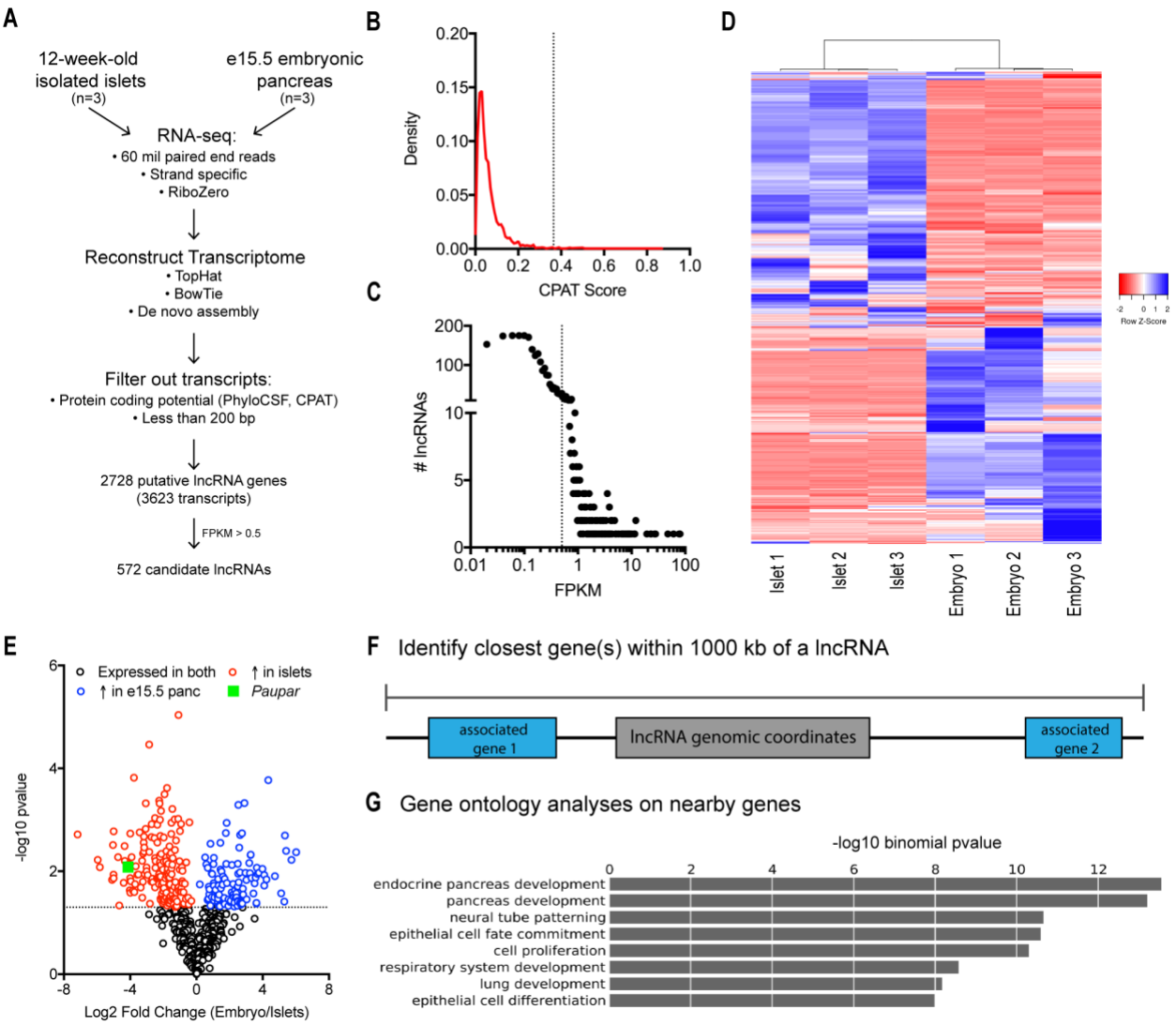


Figure 2-1. Systematic identification of developmentally regulated lncRNAs in the mouse pancreas. (A) Overview of lncRNA discovery pipeline. A total of 2728 lncRNA genes (3623 transcripts) were identified from e15.5 embryonic mouse pancreata and 12-week-old adult mouse islets. 572 lncRNAs with an FPKM > 0.5 were included in downstream analyses. (B) Histogram plot showing the number of lncRNAs corresponding to a range of Coding Potential Assessment Tool (CPAT) scores. Cutoff for inclusion in list of putative lncRNAs was CPAT score < 0.364 (shown by dotted grey line). (C) Plot showing number of lncRNAs corresponding to a range of FPKM values. Cutoff for the final list of putative lncRNAs was FPKM > 0.5 (shown by dotted

grey line). (D) Heat map of 572 candidate lncRNAs with corresponding FPKMs from e15.5 pancreas samples (n=3) and 12-week-old adult isolated islet samples (n=3). Z-scores were calculated using Pearson distance measurements. Heat map was generated with Heatmapper (http://heatmapper.ca/). (E) Volcano plot showing Log2 fold change (e15.5 pancreas/islets) and –log10 p-value for 572 candidate lncRNAs. 171 lncRNAs were upregulated in islets (red circles) and 108 lncRNAs were upregulated in embryonic pancreas (blue circles). Genes that did not meet the significance cutoff ($p > 0.05$, indicated by dotted grey line) are shown as black circles. *Paupar* lncRNA is indicated by a green square. (F) Overview of how GREAT analysis identifies set of "associated genes" within 1 MB of each lncRNA (G) GREAT outputs gene ontology analysis for set of associated genes. Each category listed is significant and shown is –log10 binomial p-value. See also Figure 2-2.

A

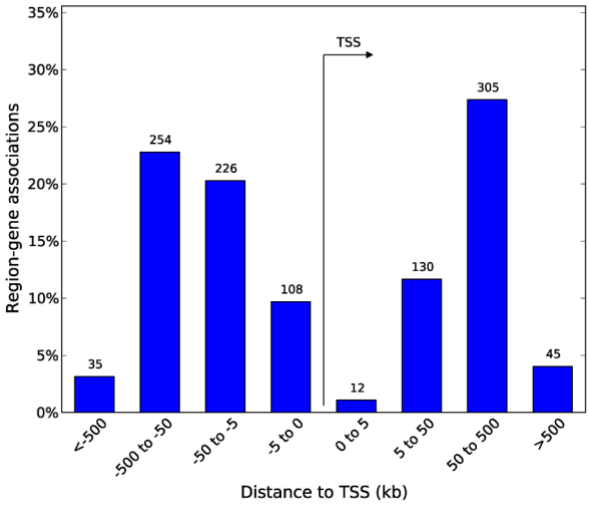


Figure 2-2. Related to Figure 2-1, Properties of pancreatic lncRNAs. (A) GREAT analysis on lncRNA candidates showing distribution of the distance between lncRNAs and the transcriptional start site of associated gene(s).

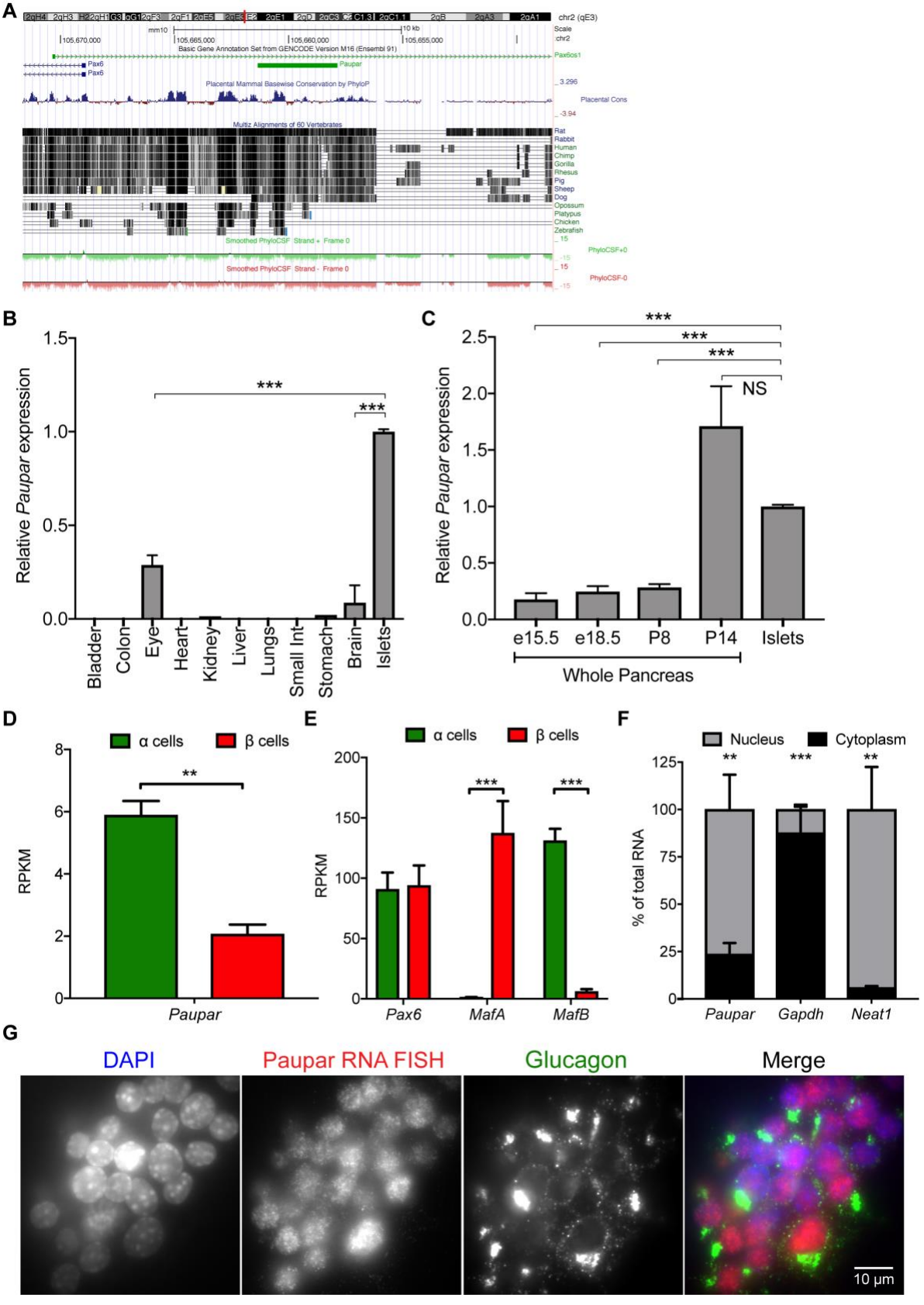
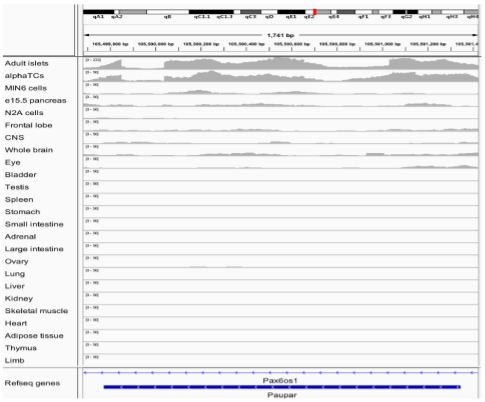
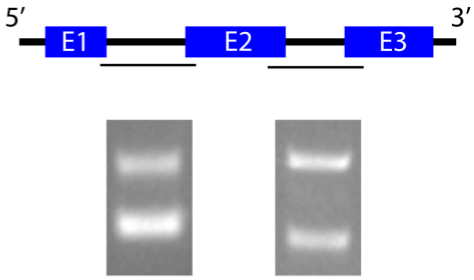

A
chr2 (qE3)
mm10
10 kb
Scale
chr2
105,670,000
105,665,000
105,660,000
105,655,000
Basic Gene Annotation Set from GENCODE Version M16 (Ensembl 91)
Pax5os1
Paupar
Pax6
Pax6
Placental Mammal Basewise Conservation by PhyloP
Placental Cons
3.296
-3.94
Multiz Alignments of 60 Vertebrates
Rat
Rabbit
Human
Chimp
Gorilla
Rhesus
Pig
Sheep
Dog
Opossum
Platypus
Chicken
Zebrafish
Smoothed PhyloCSF Strand + Frame 0
PhyloCSF+0
Smoothed PhyloCSF Strand - Frame 0
PhyloCSF-0
B
Relative Paupar expression
1.5
1.0
0.5
0.0
***
***
Bladder
Colon
Eye
Heart
Kidney
Liver
Lungs
Small Int
Stomach
Brain
Islets
C
Relative Paupar expression
2.5
2.0
1.5
1.0
0.5
0.0
***
***
***
NS
e15.5
e18.5
P8
P14
Islets
Whole Pancreas
D
α cells
β cells
RPKM
8
6
4
2
0
**
Paupar
E
α cells
β cells
RPKM
200
150
100
50
0
***
***
Pax6
MafA
MafB
F
Nucleus
Cytoplasm
% of total RNA
125
100
75
50
25
0
**
***
**
Paupar
Gapdh
Neat1
G
DAPI
Paupar RNA FISH
Glucagon
Merge
10 μm

Figure 2-3. *Paupar* is a nuclear lncRNA enriched in pancreatic alpha cells. (A) Screenshot of UCSC mm10 genome browser showing tracks for GENCODE M16, placental mammalian conservation by PhyloP, Multiz alignments, and PhyloCSF scores. (B) qRT-PCR analysis of *Paupar* RNA from 12-week-old mouse tissues, n=3. (C) qRT-PCR analysis of *Paupar* RNA from whole pancreas at indicated developmental time points and 12-week-old isolated islets n=3. (D) Plots showing *Paupar* RPKM in published datasets from adult mouse FACS purified alpha (green bar) and beta cells (red bar), n=3. (E) RPKM for *MafA, MafB* and *Pax6* based on the same published datasets used for (D) from adult mouse FACS purified alpha (green bar) and beta cells (red bar), n=3. (F) qRT-PCR analysis of *Paupar* RNA extracted from alphaTCs following cell fractionation. *Gapdh* and *Xist* are included for cell fractionation control, n=3. (G) Single molecule RNA fluorescent in situ hybridization (smFISH) images in alphaTCs showing *Paupar* (red) co-localized with antibody staining for Glucagon (green) and DAPI (blue). Scale bar indicates 10 μm. Images are representative of 3 replicate experiments. Data for (B) and others like it in the following figures are mean ± SEM, t test, NS $p > 0.05$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, and **** $p < 0.0001$. See also Figure 2-4.

A



B



C

GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGAAAACAGGAAC-
GGAGATGGGGTTTAACCTCCCTCCACCCTTGGTCTGGAAATTCCAAGGGACCCATGGACACCAGGGAGCCGAGAGAATCGCT
ATCCTGGGCCGCTCTGCCCGTCGAGGCTGCGGTGAAGATGAGGTCTGGTTGGCTTCTTTGTCCTTGGGAGCCCTCGTGGGA
ACGTTCCTTCAGCTTCCACCTAGTTGGGCCGAATCCAGGGGTGGACTTAAGCCGGAGCCCTATGGTGGATCAGTCCCATTTCT
CCCATCCTTCCTTCCACCACCAATCCTTTCTTTTCTGTTTCACAACTCCCAGGGTCTTGCTCAGCAGGATGGACCTCACTCTGG
GAGTGGGGGTGGGGGCTTCGCGAATTTGAAAGTTGTCTCTGCCTTGATCTAGTTCCAGGCTCCCCGCGGCCACTCCAAACTC
tgagaccaccagaaaagaaagaagaaaaaagtcagtatctatgcagtgcaagactctagcgctcacctttgagcgcggcctcttggaactcttaaagcccggccagggagccagg
agggggtcagtgcgcaccggcccagacaggcctctgcagcgcgcaggcgcggatccagcggccgccccctctccatgcacagggttgaggggacagaccctaagccccactctg
agcaactactgactcttaaaaaaaaaaaaagaccatctggtaattgtctttctttagatcttcccttaattatgcttttaattaaagatggttcaaggcaccggaccacatgaagggcagtaat
tacaaattacattacaaaccCGACAGATGTAGCGCGATCTGCTGGGAAAAAGGTAGCGGAATAAAAATTTGACTAAACTTTGGGATAAAT
TAATCGGCGGCCTGGAGAACAGAAGCGCAGTTCTGGTCGTGTAGACTGCAAGCAGGATTTAAGTTTTGTGTCATCATGACTGT
TTTCATATTTCTTTGACGTCTATTGATTTGGAGGCAGCTGCAAGCCTTTGTCACAGGAGAGAAGAGCAGTCACAATATATTAGCT
TGATGGCATAACTCGAACCGGAGAGGCGAGGGGAGCTGCAGTGACGGGGGCGCACGTGACCCCCTGCCCGCGGCTTGCCA
GCCCAGCCCTCACTTAAAACTTGTTCTAACTGGAACAGTAAGTATTGCTCTCCGTTCAGTCTGAAGGTGCTCTTCTGTCTAGGG
TGCGGATGGCGGTGGCAGAACCGCGGCGGCTGGGAGGCGGAAGTGTCCGAGGGTCCCGCACTACCTCTGTTGGGAGCCAC
TGAGGTTCTCGGTGGCACTTCTGGTCCGGCCTTTTGGATGAAGTTTGGGGGTTCAACTGTTTTGTCCTCAAGGCTCTGGGCC
AAGTCCTCGCTCAGTCTGGCCTTTTTGGGGGGCCTTCTCCCGAGGATGTGTGCCtgaaagacacgtggctcaggaagctattggactgcccgg
aaggtgtccccagcccccgcccctggtgcgggggctgccaaaccccctccccacgagattgtaccaggtcgggatgcaaactgcaaactcaagggagggaattaggccactct
ctttgctaagtcttacctcggcttttcttaccTCAACTCTATCCTGGTAGCACCATAGCCTGGAAGACATTTTCAAACCAGAGTCCACTCACTCTG
AGTAGCTCCGCTTTATGGAAAAACTATGTGCCCATCACTCACCCATCTATTTACAACACCGGTGCTGTGCCTCCCGGAGCATGC
AGGGACAATTAAACACGCCCACTTAGGGCGGGATGATGAAGGCTTTCTATTAACCACAGGCGTTCAGAAATGGATGCAAACCT
AATTTGGGCTTCTTGATGGGGATGGATGTTCCCAGGCAAGGTGTGTGATCCAGCCTGCTCTGGGAGAAGAGGAAGACCCAGC
GGCCTTGCATTTTCTTCTGCTTCCTTCGGAAGGAAAGTTTTCCTTTTGATTCACTGTCTCCCTGTGTCCTTCTTTTCTTTTGTATC
GCAGTTTCTGAGAGGAGTGGAGATATTTTGACTAGGCCTCCCCAGGGGCCTTTTAAATACTAGGCACCATCACCCCTGCAATAG
TTCATAGTTTCAAAATGTGGGTGTAAGGGTCTGTCTTGCTGAGGAACATCATGTAATTGGTGGTGACTGCTTAACTCATACAGTG
AAGCTATGTATTCCCTAGAGATGGCTCAAAGGCTGCTACTTTTCCAACAGTTTTTATGTGGGGGAGAAAAAAAATTTCTAACAGA
CTTTTGATATTGGGTCACAGCAAAGTCACATCTCATTGTTCACTTCCATCTCTCCCTTCTTGCTGTTGAGGCAAATACAGGAAAA
ACACAGAAATCAGGATTCAAAGCCTCTCTGGCTTTTCTACATATATGAATGAGGACTCACTGAGTCCATTGTGATACATCATGGA
AAGCAGCTCTGTTGAGACAGATATGTGGAGTTCAGAGAAATTAGAAGGACAGCAACTGTGCCAAGGGTCTCCTAGTTTTTGTG
GATAATACAGGTGATATGATGCTTTTTGTTTCAGAATGAAGCAAAAGGGGCCAGGAGGCAGACAGCCTCTCTTTTTGTGTCTGT
CTCCTTCTGCCTCTGTCTCTTTGTCTCTGTTTCTCTCTGTCTCTCTGGGTCTCTGTCTCTTGTGTGAATGTGTGTGTGTTTCCCA
GGTACCCAGGTATCCATGATTTATACTCAGAAATCATTTCTAAAAACAGAGACGTTGCCAAGTATGGTGGTACATGCTTTTAATCC
CAGCATTGGGAAAGCAGAGGTAGGCCTGTGATTCTAGAACAGCCAGGGCTACATAGTGAGACCATGTCTCAAAATAAACAAATA
GAGACAGGTCATTACAAACTTGGCTTCATCTTCTGTGGATGTCTGGGTCTTTGGAGTAGTAACGGATTAGGTCATGGTTTTTAAT
TTCTGTGACTTTCAGTGGTGGCTGTGGCCATTTTAGGGCCTAGATGTTCTCAAGAAGGAACCCGGGCTGTATTTATTGATTTCT
CCTGAAGATTTTTCTGTTAGATGAAATGTATTATTTTTAGATTAAACATCTTAGCCATCTCTAAATTCTTTTAAATTACTCTATTTGTG
ATTTTTAAAGGTGGATTTTTTTTGTTTATTTATAAGCAACTTCTAATGAATACATCACAATATCAAGAGTTTAACACAAACTTGTGT
ACAATACTGAAAAATGCTCTACAGTAGGAAAAAAAAAACAAAGAAAAACGTTAAAACTCCTATAACATTGCTTGTGATGCTAATCT
CTTCTCAGTGCAGGGAATGGAAAAATTTCAGCTGGTATTTTGTTTTCTCACTTTGTGTTTCTTCTGTAATGCTCAGCAGCACTCT
GCCTGCCTTTGCTGTCTTCCCACAGCCTGTGTGTACACTATGAATTTCAATAAATATAAATTACCGCTT

D

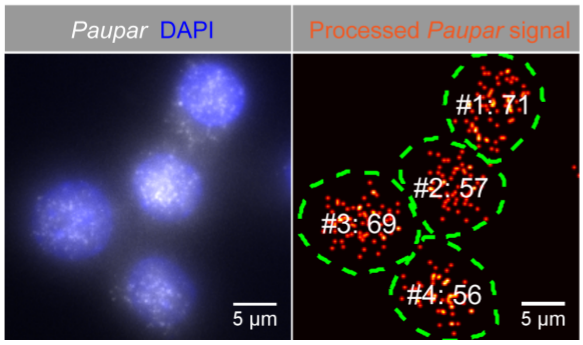


Figure 2-4. Related to Figure 2-3, Expression analysis of *Paupar* lncRNA. (A) IGV screenshot showing reads mapping to the *Paupar* locus in 25 different tissues and cell types. Y-axis maximum is set at 50 in all samples except for adult islets (top), which is set to 233 and corresponds to maximum observed *Paupar* expression (B) RT-PCR cloning of the mature *Paupar* transcript in alphaTCs using primers spanning intron/exon boundaries. (C) The full *Paupar* sequence with exons show in blue letters and introns show in lowercase black letters, mm10. (D) Single molecule FISH of *Paupar* RNA and DAPI in alphaTCs (left panel). The right panel shows processed *Paupar* signal that was then quantified according to the protocol outlined in the methods section. Scale bar indicates 5 μm.

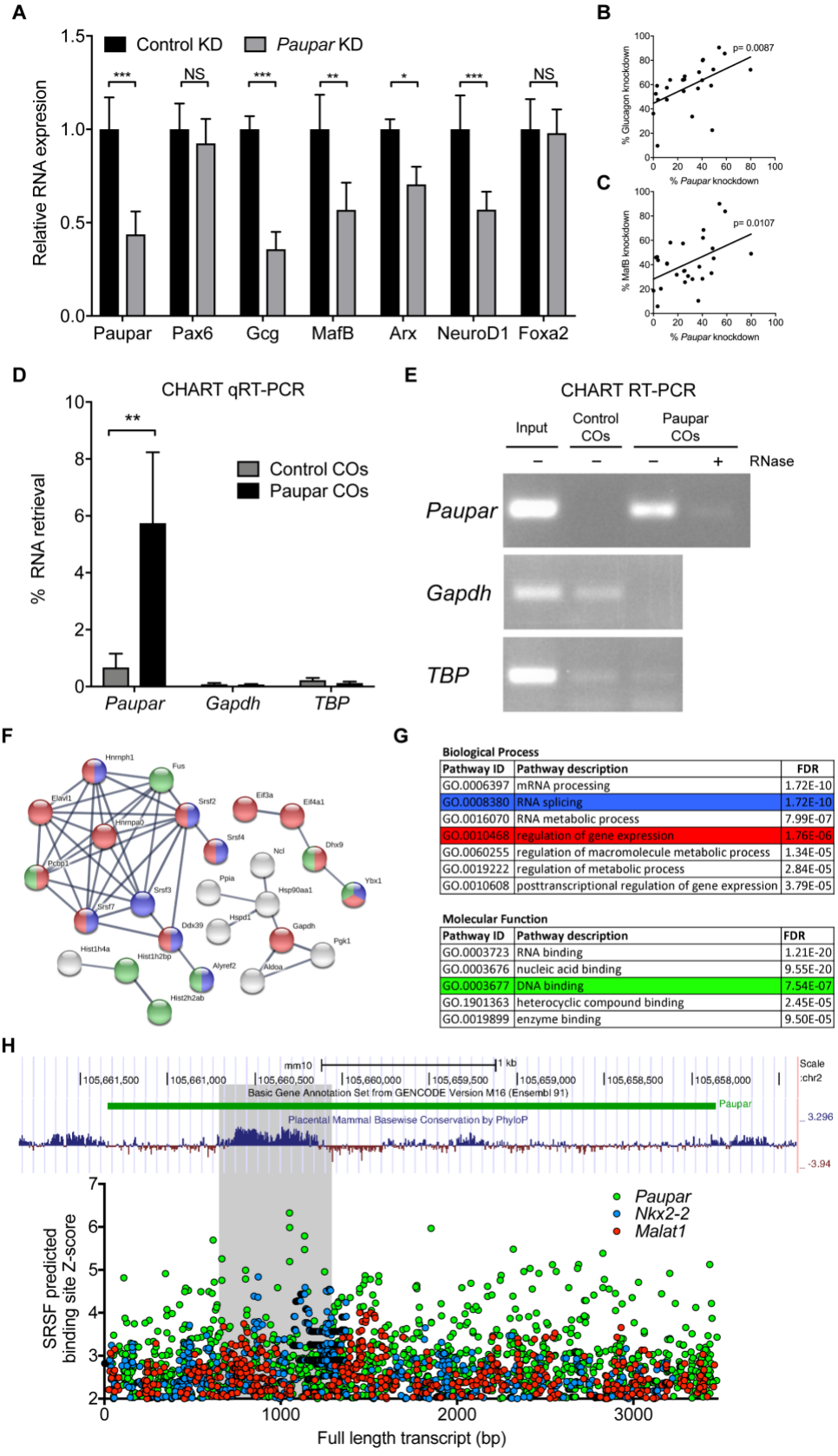

A
Control KD
Paupar KD
Relative RNA expresion
1.5
1.0
0.5
0.0
***
NS
***
**
*
***
NS
Paupar
Pax6
Gcg
MafB
Arx
NeuroD1
Foxa2
B
p= 0.0087
% Glucagon knockdown
% Paupar knockdown
C
p= 0.0107
% MafB knockdown
% Paupar knockdown
D
CHART qRT-PCR
% RNA retrieval
Control COs
Paupar COs
Paupar
Gapdh
TBP
E
CHART RT-PCR
Input
Control COs
Paupar COs
RNase
Paupar
Gapdh
TBP
F
Hnrnph1
Fus
Elavl1
Srsf2
Eif3a
Eif4a1
Hnrnpa0
Srsf4
Ncl
Dhx9
Pcbp1
Ppia
Ybx1
Srsf3
Hsp90aa1
Srsf7
Hspd1
Ddx39
Gapdh
Pgk1
Hist1h4a
Hist1h2bp
Aldoa
Alyref2
Hist2h2ab
G
Biological Process
Pathway ID | Pathway description | FDR
GO.0006397 | mRNA processing | 1.72E-10
GO.0008380 | RNA splicing | 1.72E-10
GO.0016070 | RNA metabolic process | 7.99E-07
GO.0010468 | regulation of gene expression | 1.76E-06
GO.0060255 | regulation of macromolecule metabolic process | 1.34E-05
GO.0019222 | regulation of metabolic process | 2.84E-05
GO.0010608 | posttranscriptional regulation of gene expression | 3.79E-05
Molecular Function
Pathway ID | Pathway description | FDR
GO.0003723 | RNA binding | 1.21E-20
GO.0003676 | nucleic acid binding | 9.55E-20
GO.0003677 | DNA binding | 7.54E-07
GO.1901363 | heterocyclic compound binding | 2.45E-05
GO.0019899 | enzyme binding | 9.50E-05
H
mm10
1 kb
105,661,500
105,661,000
105,660,500
105,660,000
105,659,500
105,659,000
105,658,500
105,658,000
Scale
:chr2
Basic Gene Annotation Set from GENCODE Version M16 (Ensembl 91)
Paupar
Placental Mammal Basewise Conservation by PhyloP
3.296
-3.94
SRSF predicted binding site Z-score
Paupar
Nkx2-2
Malat1
Full length transcript (bp)

Figure 2-5. *Paupar* lncRNA regulates PAX6 alpha cell target genes and interacts with several nuclear proteins involved in alternative splicing. (A) qRT-PCR analysis of *Paupar, Pax6, Glucagon (Gcg), MafB, Arx, NeuroD1,* and *Foxa2* RNA from alphaTCs following control ASOs or *Paupar* ASO KD; n=4 for each condition. (B) Linear regression plot showing significant positive correlation between percent *Paupar* KD and percent *Gcg* reduction. (C) Linear regression plot showing significant positive correlation between percent *Paupar* KD and percent *MafB* reduction. (D) Graph of qRT-PCR analysis on CHART done with control COs or *Paupar* COs showing percent RNA retrieval relative to input for *Paupar*, *Gapdh*, and *TBP*. (E) CHART RT-PCR images showing retrieval of *Paupar*, *Gapdh*, and *TBP* from pull downs with control COs, *Paupar* COs, or *Paupar* COs treated with RNase. (F) Diagram showing output from STRING analysis of *Paupar* interacting proteins found by CHART-MS. Colored nodes correspond to the categories highlighted in (G). (G) Results of GO analyses on *Paupar* interacting proteins from (G). (H) Results from RBPmap showing z-scores for predicted SRSF binding sites mapped to *Paupar* (green dots), Malat1 (red dots), and Nkx2-2 (blue dots) loci. Graph is aligned to UCSC screenshot showing the *Paupar* locus and corresponding placental mammalian basewise conservation. Grey rectangle highlights the region of *Paupar* with high sequence conservation with several highly significant SRSF binding sites. See also Figure 2-6.

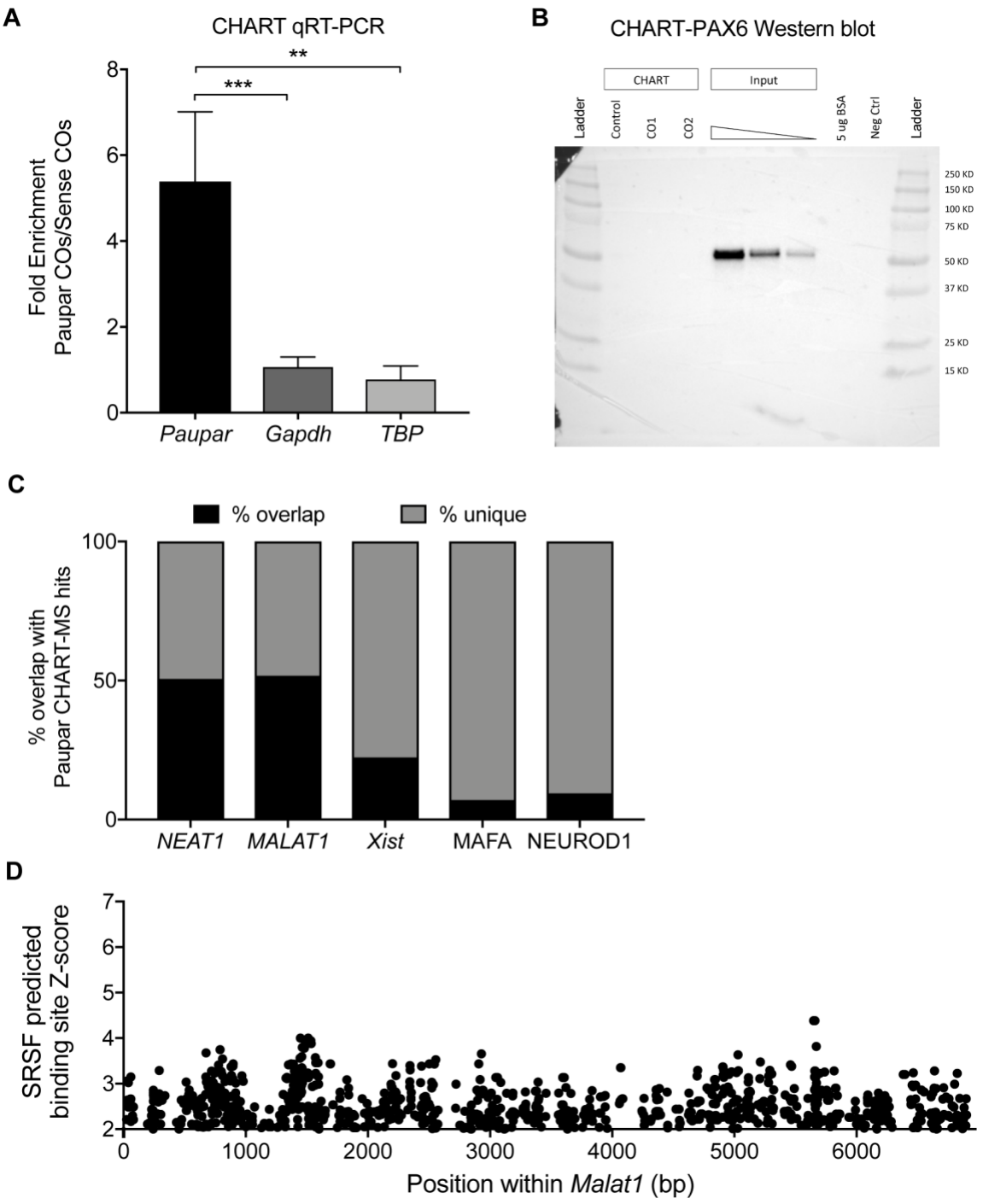


Figure 2-6. Related to Figure 2-5, Results of *Paupar* CHART-MS. (A) Graph of qRT-PCR analysis on CHART samples showing Fold Enrichment of *Paupar, Gapdh, and TBP* in experiments using *Paupar* COs compared to sense COs. (B) Representative image of a western blot for PAX6 on control CO or *Paupar* CO CHART eluate, along with varying amounts of

alphaTC input extract, 5 μg BSA, and a no protein control. (C) Plot showing percentage of unique and overlapping proteins between *Paupar*-MS and proteins retrieved by pull downs indicated on the x-axis. (D) Results from RBPmap showing z-scores for predicted SRSF binding sites mapped to the whole length *Malat1* locus.

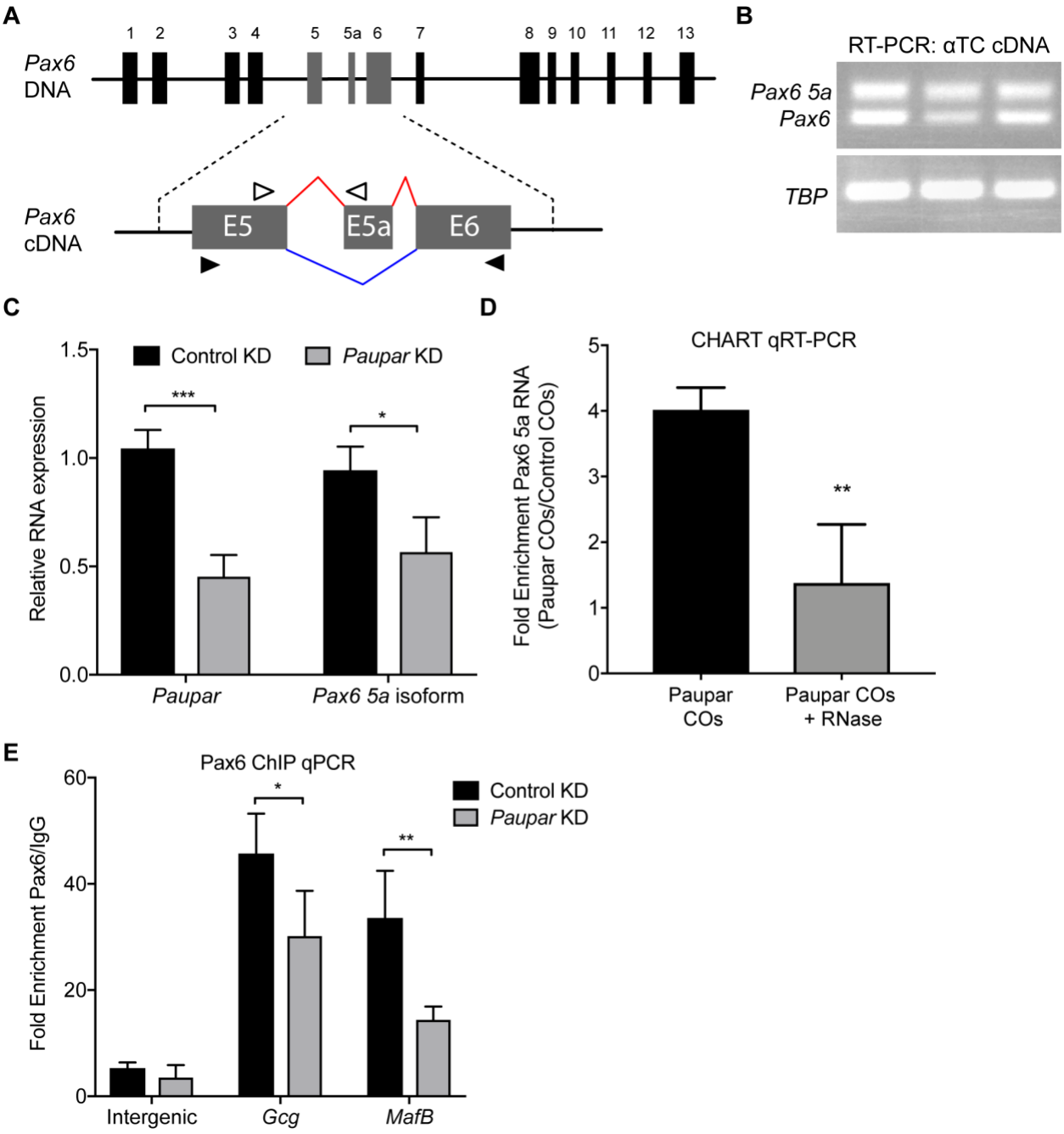


Figure 2-7. *Paupar* promotes the alternative splicing of *Pax6* to the isoform required for activation of PAX6 alpha cell target genes. (A) Schematic of the *Pax6* genomic locus with 13 constitutive exons and exon 5a. Schematic of *Pax6* cDNA with red lines showing the alternative splicing that produces the *Pax6 5a* isoform and blue lines showing the alternative splicing that produces the longer *Pax6* isoform. Black filled-in arrows are RT-PCR primers used in (B) to

amplify both isoforms. Black outlined arrows are qRT-PCR primers used in (C) and (D) to specifically amplify *Pax6 5a* RNA. (B) RT-PCR analysis showing amplification of both *Pax6* isoforms using primers indicated by black filled-in arrows shown in (A). (C) qRT-PCR analysis of *Paupar* and *Pax6* RNA from alphaTCs following control or *Paupar* KD. (D) CHART qRT-PCR showing the enrichment of *Pax6 5a* isoform in experiments using *Paupar* COs or *Paupar* COs with RNaseA treatment, compared to control COs. Primers used are shown by black outlined arrows in (A). (E) ChIP-qPCR analysis showing the fold enrichment for PAX6 antibody over IgG at intergenic, *Gcg*, and *MafB* DNA elements conducted in alphaTCs following control or *Paupar* KD.

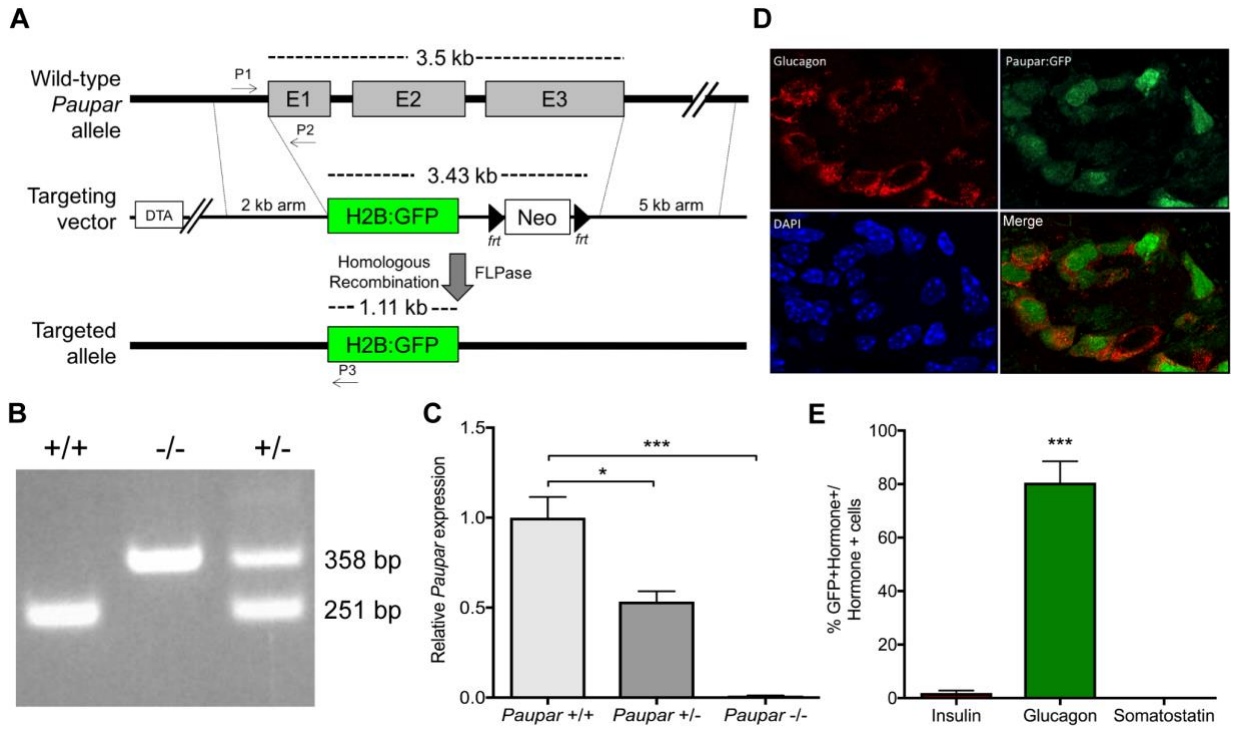


Figure 2-8. Related to Figure 2-9, Generation of *Paupar* KO mice. (A) Overview of *Paupar* knockout (KO) construct. The 3.7 kb *Paupar* locus is replaced by H2BGFP-PGKpA using BAC recombineering in mouse embryonic stem cells (mESCs). (B) Genotyping PCR strategy using P1, P2, and P3. The WT band is 251 bp and the KO band is 358 bp. (C) qRT-PCR on mouse islets showing loss of *Paupar* RNA. (D) Immunohistochemistry showing *Paupar*:H2BGFP co-localization in *Paupar* KO islets. (E) Quantification of GFP+ nuclei in insulin, glucagon, and somatostatin positive cells.

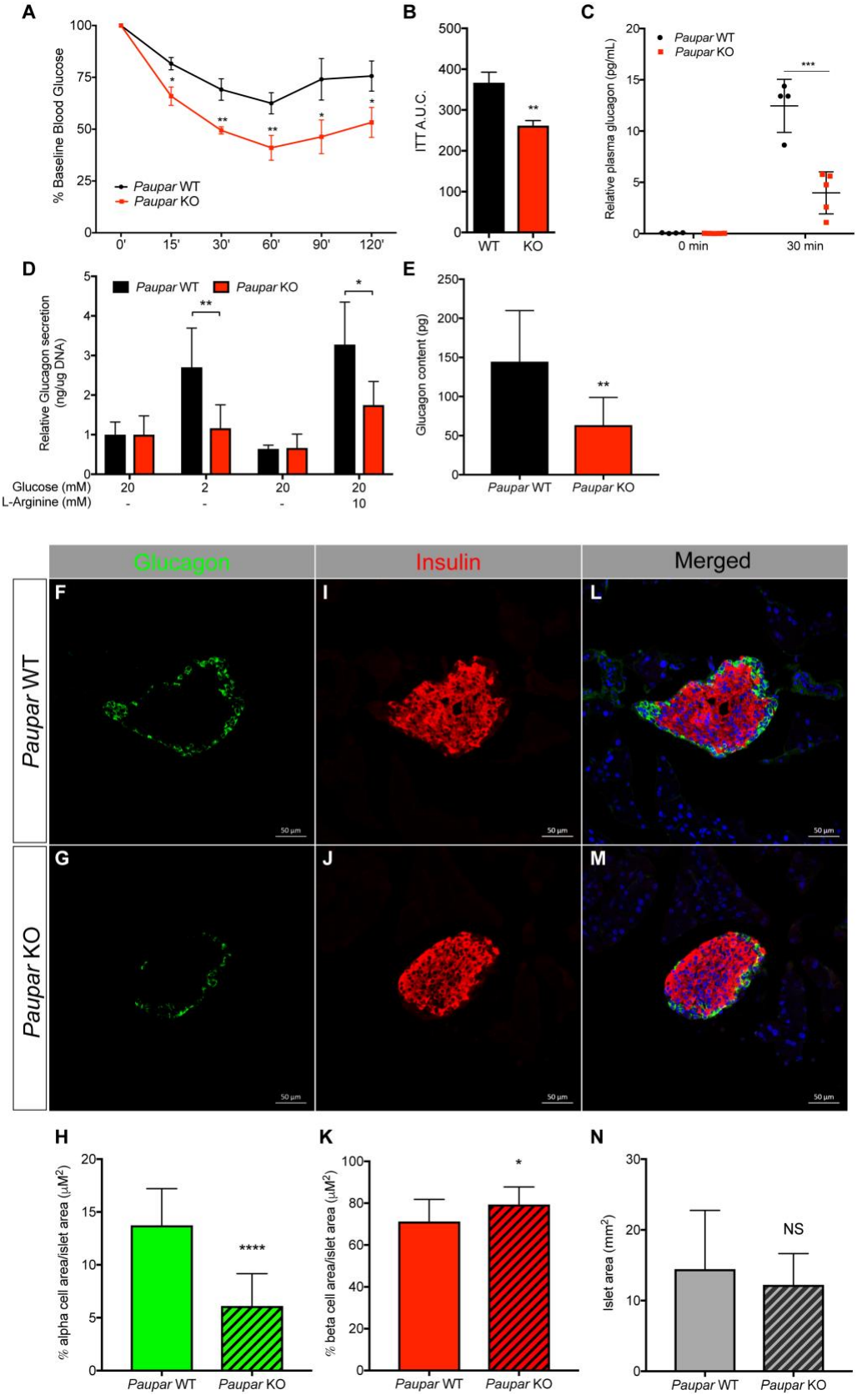

A
% Baseline Blood Glucose
Paupar WT
Paupar KO
0'
15'
30'
60'
90'
120'
B
ITT A.U.C.
WT
KO
C
Relative plasma glucagon (pg/mL)
Paupar WT
Paupar KO
0 min
30 min
D
Relative Glucagon secretion (ng/ug DNA)
Glucose (mM)
L-Arginine (mM)
E
Glucagon content (pg)
Paupar WT
Paupar KO
Glucagon
Insulin
Merged
F
I
L
G
J
M
Paupar WT
Paupar KO
50 µm
H
% alpha cell area/islet area (µM²)
****
K
% beta cell area/islet area (µM²)
N
Islet area (mm²)
NS

Figure 2-9. *Paupar* knockout mice have impaired alpha cell development and function. (A) Graph showing percent baseline blood glucose for six time points during an insulin tolerance tests on 6-week-old *Paupar* WT (black) and *Paupar* KO (red) mice. n=7-8 for each genotype. (B) Area under the curve (AUC) calculations for (A). (C) Plot showing plasma glucagon levels (pg/mL) measured by glucagon ELISA at time 0 and 30 minutes after the insulin injection during an ITT. Data points are relative values corrected for the % baseline blood glucose achieved by each animal in (A). (D) The amount of glucagon secreted by isolated islets in response to indicated stimuli. Glucagon values were measured by ELISA and are relative to μg islet DNA. Values are presented as relative to secretion during initial 20 mM glucose incubation. n=4 mice were used for each genotype and duplicate batches of 20 islets each were used for each mouse. (E) Average glucagon content measured by ELISA. n=4 mice for each genotype. (F-M) Immunofluorescence analysis of *Paupar* WT (F, I, L) and *Paupar* KO (G, J, M) pancreata showing glucagon-producing alpha cells (F, G), insulin-producing beta cells, (I, J), and merged channels with DAPI (L, M). Scale bar indicates 50 μm. Images are representative of 3 replicate experiments. (H) Quantification of immunofluorescence images showing percentage alpha cell area relative to islet area in WT (plain green bar) and KO mice (green striped bar). Data represents an average quantification of 15-20 islets from each mouse, n=3 mice per genotype. (K) Same as (H) but showing percentage beta cell area relative to islet area in *Paupar* WT (plain red bar) and *Paupar* KO mice (red striped bar). (N) Same as (H) but showing average islet in WT (plain grey bar) and KO mice (grey striped bar). See also Figure 2-10, 2-11, 2-12.

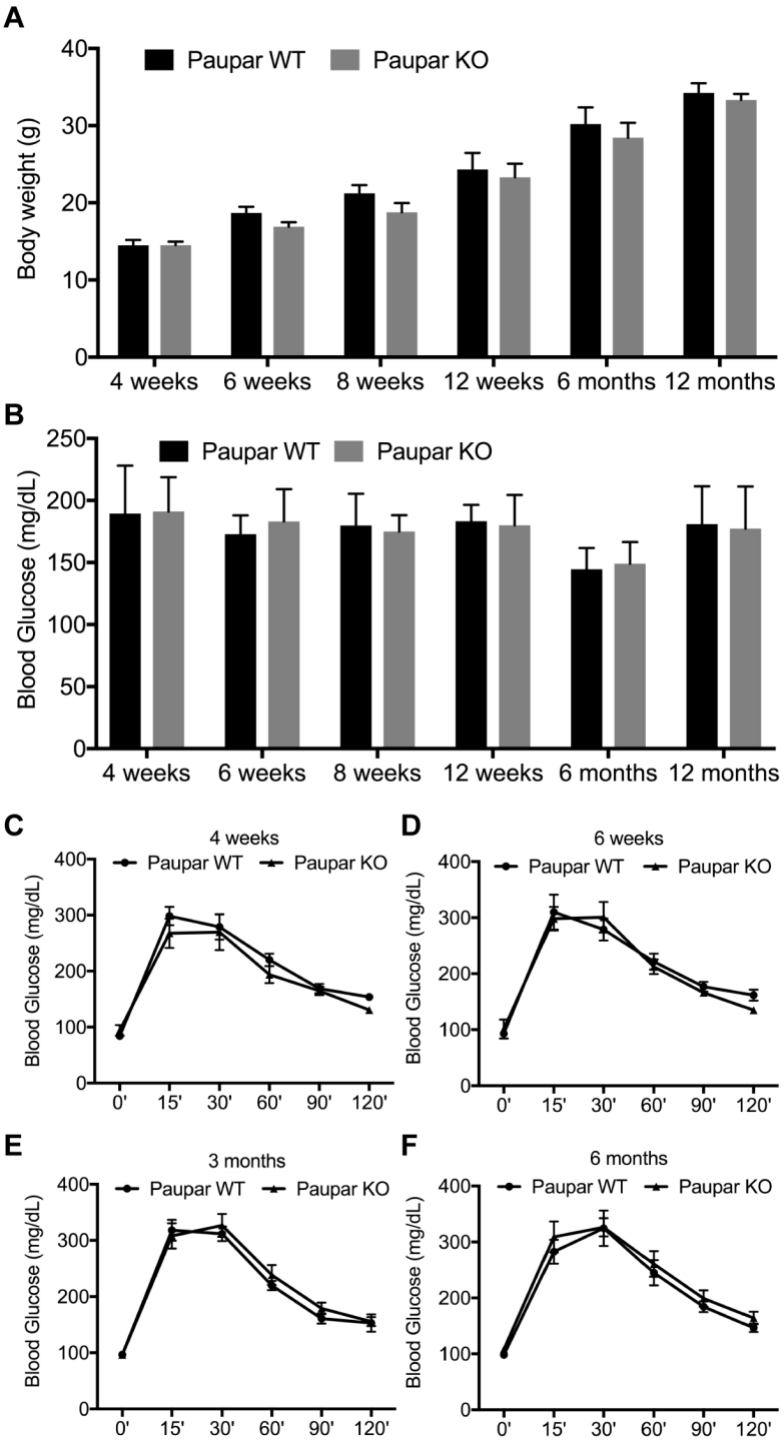


Figure 2-10. Relative to Figure 2-9, Phenotypic characterization of *Paupar* KO mice. (A) Graph showing body weight in grams for *Paupar* WT and KO mice at different ages; n=10-20 per genotype. (B) Graph showing blood glucose (in mg/dL) for the same *Paupar* WT and KO

mice in (A); n=10-20 per genotype. (C-F) Results of glucose tolerance tests on *Paupar* WT and KO mice at 4 weeks (C), 6 weeks (D), 3 months (E), and 6 months (F).

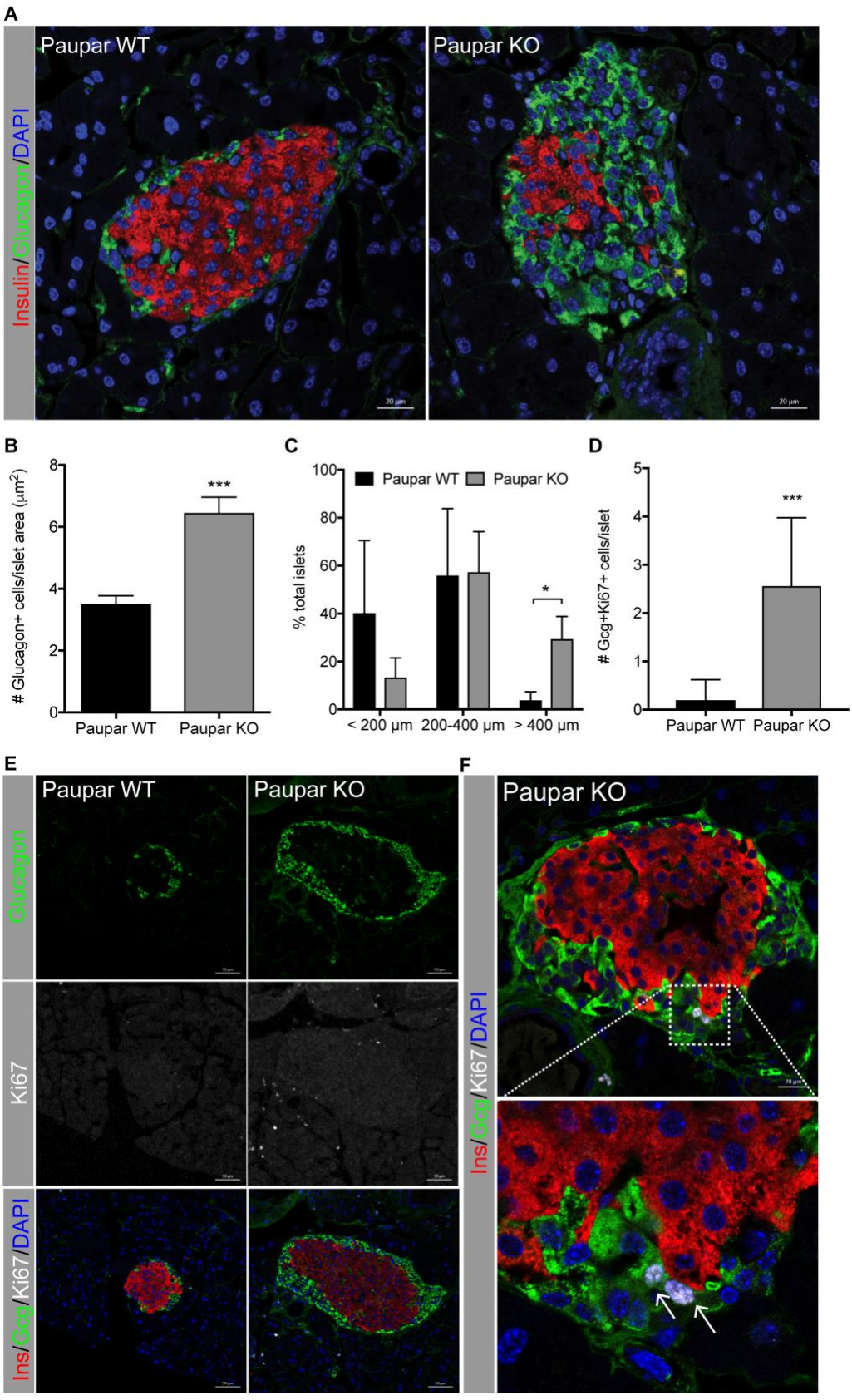

A
Paupar WT
Paupar KO
Insulin/Glucagon/DAPI
B
Glucagon+ cells/islet area (μm²)
***
Paupar WT
Paupar KO
C
% total islets
Paupar WT
Paupar KO
*
< 200 μm
200-400 μm
> 400 μm
D
Gcg+Ki67+ cells/islet
***
Paupar WT
Paupar KO
E
Paupar WT
Paupar KO
Glucagon
Ki67
Ins/Gcg/Ki67/DAPI
F
Paupar KO
Ins/Gcg/Ki67/DAPI

Figure 2-11. Relative to Figure 2-9, Aged *Paupar* KO mice develop islet and alpha cell hyperplasia. (A) Immunohistochemistry analyses of 7-month-old *Paupar* WT and KO mice showing insulin (red), glucagon (green), and DAPI (blue). (B) Morphometric analyses quantifying glucagon positive cells relative to islet area. (C) Quantification of islet size distribution in 7-month old *Paupar* WT and KO mice. (D) Quantification of Ki67 and glucagon double positive in 7-month old *Paupar* WT and KO islets. (E) Immunohistochemistry analyses of 7-month-old *Paupar* WT and KO mice showing insulin (red), glucagon (green), Ki67 (white) and DAPI (blue) (F) Magnified image of (E) showing a *Paupar* KO islet showing Ki67+ glucagon+ cells.

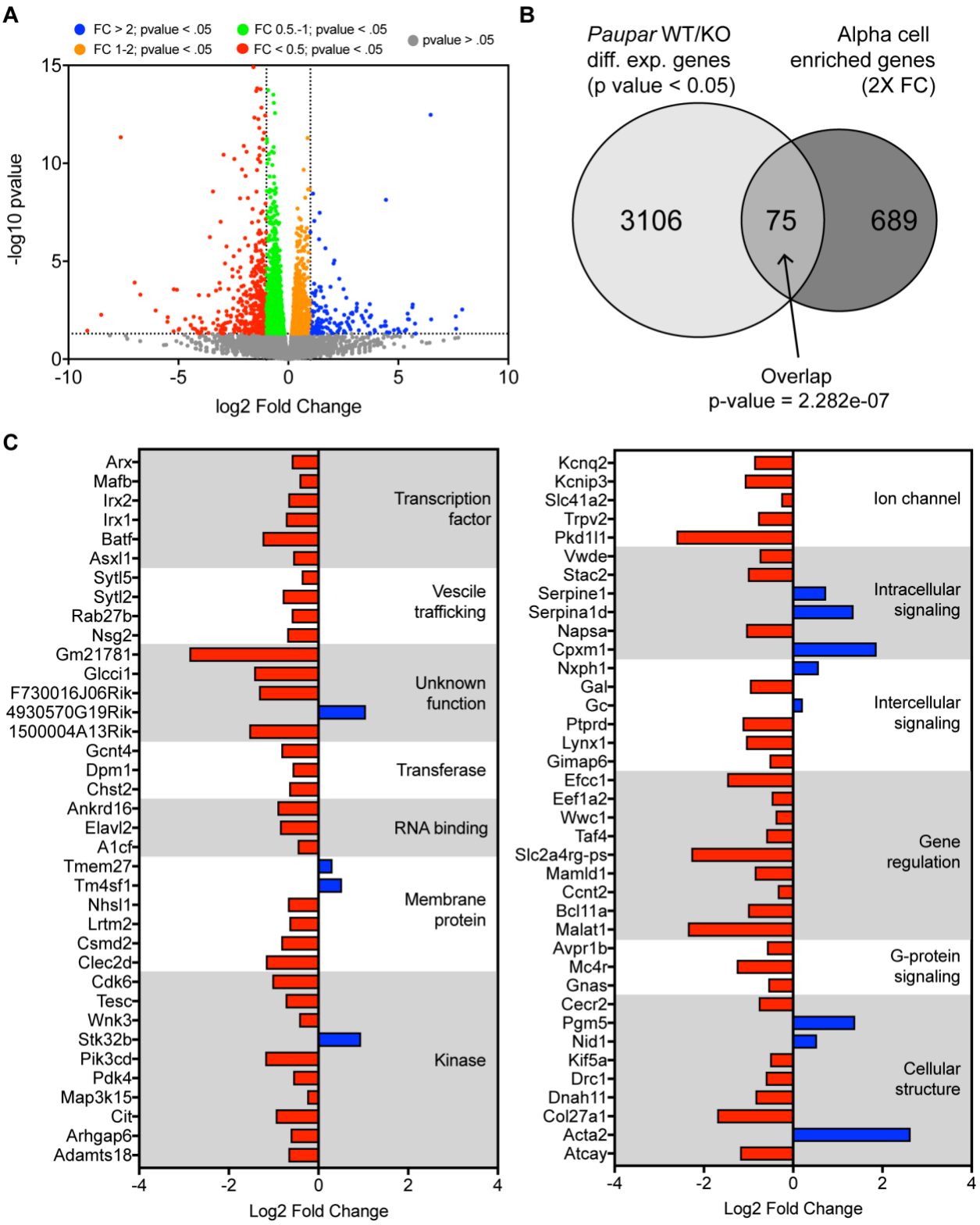

A
FC > 2; pvalue < .05
FC 1-2; pvalue < .05
FC 0.5.-1; pvalue < .05
FC < 0.5; pvalue < .05
pvalue > .05
-log10 pvalue
log2 Fold Change
B
Paupar WT/KO diff. exp. genes (p value < 0.05)
Alpha cell enriched genes (2X FC)
3106
75
689
Overlap
p-value = 2.282e-07
C
Transcription factor
Vescile trafficking
Unknown function
Transferase
RNA binding
Membrane protein
Kinase
Log2 Fold Change
Ion channel
Intracellular signaling
Intercellular signaling
Gene regulation
G-protein signaling
Cellular structure
Log2 Fold Change

Figure 2-12. *Paupar* regulates essential alpha cell genes *in vivo*. (A) Volcano plot showing Log2FC (WT/KO) and –log10 p-value highlighting 3106 differentially expressed genes (DEGs) in *Paupar* KO vs. WT mice. Stringency cutoff was p-value < 0.5 and log2FC > 2 or < 0.5, which yielded 144 upregulated genes (blue dots) and 429 genes downregulated genes (red dots). Also shown are genes with a p-value < 0.5 but with a log2FC 0.5-1 (green dots) or FC 1-2 (orange dots). Genes with a p-value > 0.5 are shown by grey dots. (B) Venn diagram showing DEGs from *Paupar* KO RNAseq compared to genes enriched > 2-fold in alpha cells compared to beta cells. Fisher exact t-test was used to quantify the significance of the overlapping 75 genes. (C) Plots showing log2 fold change values for the 75 alpha cell genes dysregulated in *Paupar* KO mice grouped by functional category.

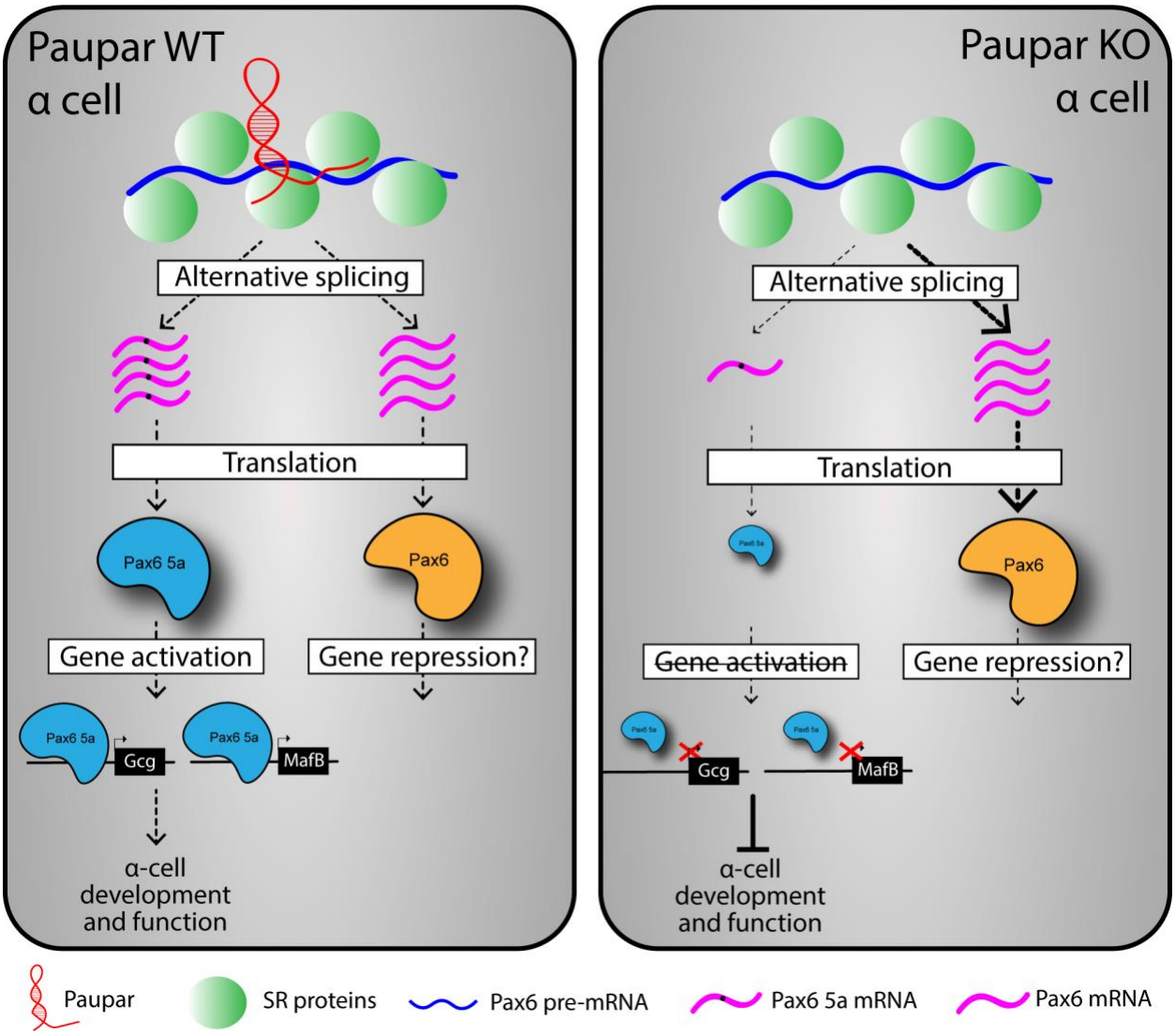


Figure 2-13. Mechanisms identified in this study through which *Paupar* regulates essential alpha cell genes. *Paupar* (red RNA molecule) is enriched in glucagon-producing alpha cells where it interacts with SR proteins (green circles) to promote the alternative splicing of the *Pax6 5a* isoform. We demonstrate that *Pax6 5a* is required for activation of the essential alpha cells genes, *Gcg* and *MafB*. Deletion of *Paupar* in vivo resulted in dysregulation of PAX6 alpha cell target genes and impaired alpha cell development and function.

Table 2-1. Detailed information on "pancreatic transcription factor associated" lncRNAs, related to Figure 2-1.

Gene ID	chr:pos (mm9)	Length	log2 FC (embryo/islets)	CPAT score	Annotation	Adjacent gene (kb to gene)	Sequence cons. (hg19)
XLOC_339243	chr2:105498019-105501500	3482	-4.121	0.208678	*Paupar* (Vance et al., 2014)	Pax6 (+8 kb)	89%
XLOC_484027	chr5:102001527-102006601	232	-7.515	0.090599	Gm20484	Nkx6-1 (86kb)	94%
XLOC_362830	chr2:79284508-79289643	671	-3.284	0.018938	Gm26755	Neurod1 (-3 kb)	63%
XLOC_240010	chr16:30008762-30018897	2309	-3.051	0.108987	4632428C04Rik	Hes1 (-40 kb)	57%
XLOC_382084	chr3:52574705-52579175	558	-2.949	0.060213	Gm2447	Foxo1 (+504 kb)	74%
XLOC_456602	chr5:29805283-29810499	335	-1.791	0.060483	None	Mnx1 (+2.8 kb)	64%
XLOC_371114	chr2:147769765-147773900	2043	0.931	0.020213	DEANR (Jiang et al., 2015)	Foxa2 (+1.5 kb)	87%
XLOC_366562	chr2:115891833-115892847	584	1.090	0.037845	2700033N17Rik	Meis2 (+1 kb)	87%
XLOC_624540	chr9:74697348-74701625	4173	8.437	0.012125	Gm16551	Hnf6 (-8 kb)	95%

Table 2-2. Proteins enriched by Capture Hybridization Analysis of RNA Targets using *Paupar* Capture Oligos (COs) or Control COs, related to Figure 2-5.

Protein	Gene	Control COs_1	Control COs_2	Paupar COs_1	Paupar COs_2
HNRH1	Hnrnph1	0	1	50	40
EF1A1	Eef1a1	1	0	21	15
SRSF3	Srsf3	0	0	9	7
DDX	Ddx17	0	0	10	5
G3P	Gapdh	1	1	15	13
GLUC	Gcg	0	0	7	7
HS90A	Hsp90aa1	0	0	9	3
ROA0	Hnrnpa0	7	6	44	43
H4	Hist1h4a	5	6	37	37
1433	Ywhab	0	0	8	3
FUS	Fus	0	0	6	5
IF4A1	Eif4a1	0	0	9	2
RBM14	Rbm14	1	0	13	3
FUBP1	Fubp1	0	0	5	5
PPIA	Ppia	0	0	6	4
YBOX1	Ybx1	0	0	8	2
H2A	Hist2h2ab	5	5	31	23
CLH1	Cltc	0	0	6	3
HMGB1	Hmgb1	0	0	6	3
NUCL	Ncl	0	0	6	3
PYC	Pc	0	0	7	2
H1	Hist1h1a	5	8	30	34
ALDOA	Aldoa	0	0	5	3
MCM2	Mcm2	0	0	6	2
SRSF7	Srsf7	0	0	4	4
HMGN1	Hmgn1	1	2	16	2
MATR3	Matr3	0	0	4	3
PRDX1	Prdx1	0	0	3	4
SFPQ	Sfpq	0	1	4	6
ALRF2	Alyref2	0	0	3	3
NDKA	Nme1	0	0	3	3
SRRM1	Srrm1	0	0	4	2
SRSF2	Srsf2	0	0	3	3
ACLY	Acly	0	0	3	2
CH60	Hspd1	0	0	5	0
DHX9	Dhx9	0	0	2	3
DX39A	Ddx39a	0	0	2	3
GSTP1	Gstp1	0	0	2	3
MYEF2	Myef2	0	0	3	2
PCBP1	Pcbp1	0	0	4	1

PGK1	Pgk1	0	0	3	2
PUR6	Paics	0	0	3	2
STMN1	Stmn1	0	0	2	3
TADBP	Tardbp	0	0	1	4
H2B	Hist1h2bp	15	19	36	38
EIF3A	Eif3a	0	0	4	0
ELAV1	Elavl1	0	0	3	1
LC7L2	Luc7l2	0	0	3	1
NUP93	Nup93	0	0	2	2
SARNP	Sarnp	0	0	2	2
SF3B1	Sf3b1	0	0	3	1
SRSF4	Srsf4	0	0	3	1
SYDC	Dars	0	0	2	2
TCEA1	Tcea1	0	0	4	0
TIF1B	Trim28	0	0	1	3
H3	H3	4	6	11	10

Table 2-3. The misexpression of canonical alpha cell transcription factors in *Paupar* KO vs. *Paupar* WT islets, related to Figure 2-12.

Genes	baseMean *Paupar* WT	baseMean *Paupar* KO	Fold Change (KO/WT)	p-value
Pax6	6363.059389	6632.967471	1.042417	0.651773473
Arx	599.440673	395.488637	0.659776	1.10E-05
Mafb	2497.636452	1858.399089	0.744067	0.000406464
Irx1	474.594955	286.584016	0.603858	0.000533321
Irx2	567.904993	355.655694	0.626272	5.36E-05

Table 2-4. Oligonucleotides used in Chapter 2.

Generation of *Paupar* KO mouse			
Name	Sequence	Source	Comment
H2BGFP sequencing primer_FWD	CAACAAGCGCTCGACCATCAC	This study	
H2BGFP sequencing primer_REV	AGTCGATGCCCTTCAGCTCGAT	This study	
Amplify H2BGFP_overlap pL451_FWD	gtcgacggtatcgataATGCCAGAGCCAGCGAAG	This study	
Amplify H2BGFP_overlap for pA-_REV	gacttacagTTACTTGTACAGCTCGTCCATG	This study	
Amplify pA_overlap H2BGFP_FWD	gtacaagtaaCTGTAAGTCTGCAGAAATTG	This study	
Amplify pA_ overlap pL451_REV	cttcggaattcgatatcaCAGCTTCTGATGGAATTAGAAC	This study	
Add 60 bp Paupar homology arms on H2BGFP-pA-FNF_FWD	AAACAGGAACGGAGATGGGGTTTAACCTCCCTCCACCCTTGGTCTGGAAATTCCAAGGGACAATGCCAGAGCCAGCGAAG	This study	
Add 60 bp Paupar homology arms on H2BGFP-pA-FNF_REV	CTTTGAGCCGTCTGAAGAGCGAGAACAATGAGACATGAAGTAACTACTCTTTACTATTTTTATTATGTACCTGACTGATG	This study	
H2BGFPpA Paupar KI 5' screening_FWD	TACACGGAGGCCGATTTTCC	This study	
H2BGFPpA Paupar KI 5' screening_REV	ATGGTCGAGCGCTTGTTGTA	This study	
H2BGFPpA Paupar KI 3' screening_FWD	GCATCGCCTTCTATCGCCTT	This study	
H2BGFPpA Paupar KI 3' screening_REV	GCTCGCGCTCCTCTAGATTG	This study	
LARH primer for DTA	GGGCTTGGTCATGACTGGTTGTTCCGGTAGTGTTT	This study	

retrieval of H2BGFP from BAC	GCATTGTGGCCTCGACAGGCTGTCTGCGGCCGCCAATTCGCCCTATAGTGAGT		
RAFH primer for DTA retrieval of H2BGFP from BAC	GGAACAGCTACTGAGGCAGGCAGATGACCAATGAAGGATAGTACATAAGCCACCAGTGAGGGCGCGCCCAGTGTGGTTTTCAAGAGGA	This study	
H2BGFPpAFNF with 2kb and 5kb arms in pMCS-DTA 5' screening_FWD	TTGGGTAACGCCAGGGTTTT	This study	
H2BGFPpAFNF with 2kb and 5kb arms in pMCS-DTA 5' screening_REV	TGTTAATGGCTGTGGTCGCA	This study	
H2BGFPpAFNF with 2kb and 5kb arms in pMCS-DTA 3' screening_FWD	GGCCTGAGAGTTCTCCGCTA	This study	
H2BGFPpAFNF with 2kb and 5kb arms in pMCS-DTA 3' screening_REV	GCAAAACCACACTGCTCGAC	This study	
5' screening of Target vector insertion into Paupar locus_FWD	GACCTCTGTACCTGAAGTTTGACAT	This study	
5' screening of Target vector insertion into Paupar locus_REV	TCAGAACCTTGTACACATAGATGGA	This study	
Genotyping			
P1_FWD primer 5' upstream of Paupar	CTCCGTCTTTGGAATCCTTG	This study	Figures 2-8A and 2-8B
P2_REV primer in WT Paupar	CCAGGATAGCGATTCTCTCG	This study	Figures 2-8A and 2-8B

P3_REV primer in H2BGFP	CTGCTTCAGAACCTTGTACACAT	This study	Figures 2-8A and 2-8B
FLPe_FWD	CACTGATATTGTAAGTAGTTTGC	This study	
FLPe_REV	CTAGTGCGAAGTAGTGATCAGG	This study	
NeoR_FWD	CGATGATCTCGTCGTGACCC	This study	FWD primer in Neo of Paupar KI to determine Neo has been FLPed out
PauparNeo_REV	CTTTGAGCCGTCTGAAGAGC	This study	REV primer in Paupar 3' gene to determine Neo has been FLPed out
smFISH Probes			
Paupar_smFISH_1	GTGACTGCTCATATTCAAAG	This study	Probe was labeled with Alexa647 at the 3' end and purchased from IDT
Paupar_smFISH_2	CAAAGACGGAGAATGGAAAT	This study	Probe was labeled with Alexa647 at the 3' end and purchased from IDT
Paupar_smFISH_3	TGTATCAGCACCGCGTTGCC	This study	Probe was labeled with Alexa647 at the 3' end and purchased from IDT
Paupar_smFISH_4	ACGCATAAATAATGCAGACG	This study	Probe was labeled with Alexa647 at the 3' end and purchased from IDT
Paupar_smFISH_5	ATAGCGATTCTCTCGGCTCC	This study	Probe was labeled with Alexa647 at the 3' end and purchased from IDT
Paupar_smFISH_6	ATCTTCACCGCAGCCTCGAC	This study	Probe was labeled with Alexa647 at the 3' end and purchased from IDT
Paupar_smFISH_7	ACAAAGAAGCCAACCAGACC	This study	Probe was labeled with Alexa647 at the 3' end and purchased from IDT
Paupar_smFISH_8	TGGAAGCTGAAGGAACGTTC	This study	Probe was labeled with Alexa647 at the 3' end and purchased from IDT
Paupar_smFISH_9	GGTGGTGGAAGGAAGGATGG	This study	Probe was labeled with Alexa647 at the 3' end and purchased from IDT
Paupar_smFISH_10	CCATCCTGCTGAGCAAGACC	This study	Probe was labeled with Alexa647 at the 3' end and purchased from IDT
Paupar_smFISH_11	GGCCGCCGATTAATTTATCC	This study	Probe was labeled with Alexa647 at the 3' end and purchased from IDT

Paupar_smFISH_12	ACGACCAGAACTGCGCTTCT	This study	Probe was labeled with Alexa647 at the 3' end and purchased from IDT
Paupar_smFISH_13	TTAAATCCTGCTTGCAGTCT	This study	Probe was labeled with Alexa647 at the 3' end and purchased from IDT
Paupar_smFISH_14	CTCCAAATCAATAGACGTCA	This study	Probe was labeled with Alexa647 at the 3' end and purchased from IDT
Paupar_smFISH_15	TGTGACAAAGGCTTGCAGCT	This study	Probe was labeled with Alexa647 at the 3' end and purchased from IDT
Paupar_smFISH_16	TATTGTGACTGCTCTTCTCT	This study	Probe was labeled with Alexa647 at the 3' end and purchased from IDT
Paupar_smFISH_17	GAGTTATGCCATCAAGCTAA	This study	Probe was labeled with Alexa647 at the 3' end and purchased from IDT
Paupar_smFISH_18	ACTGTTCCAGTTAGAACAAG	This study	Probe was labeled with Alexa647 at the 3' end and purchased from IDT
Paupar_smFISH_19	GACTGAACGGAGAGCAATAC	This study	Probe was labeled with Alexa647 at the 3' end and purchased from IDT
Paupar_smFISH_20	CCTAGACAGAAGAGCACCTT	This study	Probe was labeled with Alexa647 at the 3' end and purchased from IDT
RNAi (Antisense oligos)			
Paupar_ASO_1	mG*mU*mG*mA*mG*G*T*C*C*A*T*C*C*T*G*mC*mU*mG*mA*mG	This study	ASOs were modified and ordered from IDT as follows (*= Phosphorothioated DNA bases; m*= Phosphorothioated 2' O-methyl bases)
Paupar_ASO_2	mU*mA*mU*mU*mG*T*G*A*C*T*G*C*T*C*T*mU*mC*mU*mC*mU	This study	ASOs were modified and ordered from IDT as follows (*= Phosphorothioated DNA bases; m*= Phosphorothioated 2' O-methyl bases)
Paupar_ASO_3	mA*mC*mG*mA*mC*C*A*G*A*A*C*T*G*C*G*mC*mU*mU*mC*mU	This study	ASOs were modified and ordered from IDT as follows (*= Phosphorothioated DNA bases; m*= Phosphorothioated 2' O-methyl bases)

Scrambled_ASO	mU*mC*mG*mU*mA*A*C*A*C*G*T*C*T*A*T*mA*mC*mG*mC*mU	This study	ASOs were modified and ordered from IDT as follows (*= Phosphorothioated DNA bases; m*= Phosphorothioated 2' O-methyl bases)
qRT-PCR			
Paupar_qRTPCR_FWD	AAATTAATCGGCGGCCTGGA	This study	
Paupar_qRTPCR_REV	CTCCGGTTCGAGTTATGCCA	This study	
TBP_qRTPCR_FWD	ACCCTTCACCAATGACTCCTATG	This study	
TBP_qRTPCR_REV	ATGATGACTGCAGCAAATCGC	This study	
Gapdh_qRTPCR_FWD	TGCCCCCATGTTTGTGATG	This study	
Gapdh_qRTPCR_REV	TGTGGTCATGAGCCCTTCC	This study	
Neat1_qRTPCR_FWD	TGGCCCCTTTTGTTCATTAGC	This study	
Neat1_qRTPCR_REV	TGGAAGGCCATTGTTTCAGG	This study	
Pax6_5a_qRTPCR_FWD	CGGAGTGAATCAGCTTGGTG	This study	Figures 2-7A, C, and D (white arrows)
Pax6_5a_qRTPCR_FWD	TTTCATTGTCCAGCACCTGG	This study	Figures 2-7A, C, and D (white arrows)
RT-PCR			
Pax6_RTPCR_FWD	CGGAGTGAATCAGCTTGGTG	This study	Figures 2-7A and 2-7B (black arrows)
Pax6_RTPCR_REV	GTCTGATGGAGCCAGTCTCG	This study	Figures 2-7A and 2-7B (black arrows)
Paupar_Exon 1_FWD	TGATCTAGTTCCAGGCTCCCC	This study	Figures 2-4A and 2-4B
Paupar_Exon 2_REV	CTCCCAACAGAGGTAGTGCG	This study	Figures 2-4A and 2-4B
Paupar_Exon 1_FWD	TGATCTAGTTCCAGGCTCCCC	This study	Figures 2-4A and 2-4B
Paupar_Exon 2_REV	CTCCCAACAGAGGTAGTGCG	This study	Figures 2-4A and 2-4B
CHART capture oligos (COs)			
Paupar_CHART_1	GGCCGCCGATTAATTTATCC	This study	Capture oligos were modified at the 3' end with "/iSp18//3BioTEG/" and ordered from IDT
Paupar_CHART	ACGACCAGAACTGCGCT	This study	Capture oligos were

_2	TCT		modified at the 3' end with "/iSp18//3BioTEG/" and ordered from IDT
Paupar_CHART_3	TTAAATCCTGCTTGCAGTCT	This study	Capture oligos were modified at the 3' end with "/iSp18//3BioTEG/" and ordered from IDT
Paupar_CHART_4	CTCCAAATCAATAGACGTCA	This study	Capture oligos were modified at the 3' end with "/iSp18//3BioTEG/" and ordered from IDT
Paupar_CHART_5	TGTGACAAAGGCTTGCAGCT	This study	Capture oligos were modified at the 3' end with "/iSp18//3BioTEG/" and ordered from IDT
Paupar_CHART_6	TATTGTGACTGCTCTTCTCT	This study	Capture oligos were modified at the 3' end with "/iSp18//3BioTEG/" and ordered from IDT
Paupar_CHART_7	GAGTTATGCCATCAAGCTAA	This study	Capture oligos were modified at the 3' end with "/iSp18//3BioTEG/" and ordered from IDT
Paupar_CHART_8	ACTGTTCCAGTTAGAACAAG	This study	Capture oligos were modified at the 3' end with "/iSp18//3BioTEG/" and ordered from IDT
Paupar_CHART_9	GACTGAACGGAGAGCAATAC	This study	Capture oligos were modified at the 3' end with "/iSp18//3BioTEG/" and ordered from IDT
Paupar_CHART_10	CCTAGACAGAAGAGCACCTT	This study	Capture oligos were modified at the 3' end with "/iSp18//3BioTEG/" and ordered from IDT
Paupar_Sense_CHART	GGATAAATTAATCGGCGGCC	This study	Capture oligos were modified at the 3' end with "/iSp18//3BioTEG/" and ordered from IDT
Scrambled_CHART	TCGTAACACGTCTATACGCT	This study	Capture oligos were modified at the 3' end with "/iSp18//3BioTEG/" and ordered from IDT
ChIP-qPCR			
Negative control ChIP FWD	GGCTTGGGTTGAGGCTGGA	Keller et al 2007	Negative control for alphaTC cells

Negative control ChIP REV	CGGAGTGGCGGCGATAGAAG	Keller et al 2007	Negative control for alphaTC cells
Gcg ChIP FWD	AAGCAGATGAGCAAAGTGAGTG	Schaffer et al 2013	Gcg promoter
Gcg ChIP REV	AGGCTGTTTAGCCTTGCAGATA	Schaffer et al 2013	Gcg promoter
MafB ChIP FWD	ACTTGGGGTCGCACTTTATG	Menéndez-Gutiérrez et al. 2015	MafB promoter
MafB ChIP REV	TCTGTGCACTCAGTGGCTCT	Menéndez-Gutiérrez et al. 2015	MafB promoter

Table 2-5. Key Resources used in Chapter 2.

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
DAPI (4′,6-diamidino-2-phenylindole)	Thermo Fisher	Cat#D1306; RRID:AB_2629482
Rabbit monoclonal anti-Glucagon	Santa Cruz	Cat#sc-514592; RRID:AB_2629431
Guinea pig anti-Insulin	DAKO	Cat#A0564; RRID:AB_2617169
Rat monoclonal anti-Somatostatin	Abcam	Cat#AB30788; RRID:AB_778010
Rabbit polyclonal anti-Ki67	EMD Millipore	Cat#AB9260; RRID:AB_2142366
Alexa Fluor 488 anti-rabbit	Jackson ImmunoResearch	Cat#711-545-152; RRID:AB_2313584
Cy3 anti-guinea pig	Jackson ImmunoResearch	Cat#706-165-148; RRID:AB_2340460
Alexa Fluor 647 anti-rabbit	Jackson ImmunoResearch	Cat#711-605-152; RRID:AB_2492288
Alexa Fluor 647 anti-rat	Jackson ImmunoResarch	Cat#712-605-153; RRID:AB_2340694
Purified rabbit anti-Pax6 for ChIP	Biolegend	Cat#901302; RRID:AB_2749901
Rabbit IgG for ChIP	Millipore Sigma	Cat#I5006; RRID:AB_1163659
Bacterial and Virus Strains		
One Shot TOP10 Chemically Competent *E. coli*	Invitrogen	Cat#3879S
SW106 cells	Warming et al., 2005	N/A
Biological Samples		
Mouse embryonic stem cells (mESCs)	129SV background	N/A
Chemicals, Peptides, and Recombinant Proteins		
Proteinase K	Promega	Cat#MC5005
Q5 HiFi DNA polymerase	NEB	Cat#M0491S
Gibson assembly kit	NEB	Cat#E2611S
Terrific Broth (TB)	Difco	Cat#243820
LB Agar	Difco	Cat#240110
RNase H	NEB	Cat#M0297S
DpnI	NEB	Cat#R0176S
EcoRI-HF	NEB	Cat#R3101S
HindIII-HF	NEB	Cat#R3104S
Go Taq DNA Polymerase	Promega	Cat#M3001
dNTPs nucleotide mix	Roche	Cat#11581295001
Sodium borohydride	Millipore Sigma	Cat#452882

Formamide	Roche	Cat#11814320001
Sucrose	Millipore Sigma	Cat#S0389
DMEM media	Thermo Fisher	Cat#11995
Penicillin-Streptomycin	Thermo Fisher	Cat#15140163
HEPES buffer	Gibco	Cat#15630-080
Non-essential amino acids solutions	Sigma	Cat#M7145
Bovine Serum Albumin	Thermo Fisher	Cat#15260037
Sodium bicarbonate	Thermo Fisher	Cat#MT25035CI
Antibiotic-Antimycotic	Thermo Fisher	Cat#15240062
M199 media	Invitrogen	Cat#11150067
Fetal Bovine Serum	Gemini Bio Products	Cat#100106
Precision plus protein kaleidoscope protein standard	Bio-rad	Cat#1610375
Lipofectamine 2000 transfection Reagent	Thermo Fisher	Cat#11668-019
D-Glucose	Millipore Sigma	Cat#G8270
L-Arginine	Millipore Sigma	Cat#A5006
Insulin (NovoLog)	Novo Nordisk	U-100
iQ Sybr Green Supermix	Bio-rad	Cat#1708880
Real Time PCR Mastermix for Taqman	Eurogentec	Cat#RTQP2X0315+
iTaq Universal SYBR Green One-Step Kit	Bio-rad	Cat#1725150
Collagenase P	Millipore Sigma	Cat#11213857001
Donkey Serum	Sigma	Cat#D9663
Dynabeads MyOne Streptavidin C1	Invitrogen	Cat#65001
Paraformaldehyde EM Grade	Polysciences, Inc.	Cat#00380
Trizol LS	Invitrogen	Cat#10296010
Critical Commercial Assays		
RNeasy Plus Mini Kit	Qiagen	Cat#74134
RNeasy Plus Micro Kit	Qiagen	Cat#74034
TruSeq Stranded Total RNA (with Ribo-Zero)	Illumina	Cat#RS-122-2201
Pierce BCA protein assay kit	Thermo Fisher	Cat#23225
SuperScript III First-Strand Synthesis System	Invitrogen	Cat#18080051
Protein and RNA Isolation System (PARIS) Kit	Ambion	Cat#AM1921
Glucagon ELISA	Mercodia	Cat#10-1281-01
ChIP-IT High Sensitivity kit	Active Motif	Cat#53040
Deposited Data		

Raw and processed e15.5 embryonic mouse pancreas RNA-sequencing data	This study	GEO: GSE122033
Raw and processed 12-week-old adult mouse islets RNA-sequencing data	This study	GEO: GSE122033
Raw and processed 6-week-old Paupar WT and KO RNA-sequencing data	This study	GEO: GSE121884
ENCODE RNA-sequencing data	ENCODE Project Consortium. 2012	PRJNA66167
RNA-seq FACS purified mouse alpha and beta cells	DiGruccio et al., 2015	GEO: GSE80673
RNA-seq FACS purified human alpha and beta cells	Ackermann et al., 2016	GEO: GSE76268
Experimental Models: Cell Lines		
alpha TC1 clone 6 (αTC) cells	American Type Culture Collection	CRL-2934
Mouse Insulinoma (MIN6) cells	Miyazaki et al., 1990	N/A
Experimental Models: Organisms/Strains		
Paupar$^{tm(H2BGFP)Suss}$	This study	N/A
C57BL/6J	Jackson Laboratories	Cat#000664
FLPe transgenic mice	Jackson Laboratories	Cat#003946
Oligonucleotides		
Genotyping primers	This study	Table 2-4
smFISH probes	This study (IDT)	Table 2-4
Antisense Oligonucleotide (ASO) sequences	This study (IDT)	Table 2-4
qRT-PCR primers	This study	Table 2-4
CHART capture oligonucleotides (COs)	This study (IDT)	Table 2-4
ChIP-qPCR primers	Keller et al 2007; Schaffer et al 2013; Menéndez-Gutiérrez et al. 2015	Table 2-4
Pax6 TaqMan AOD	Thermo Fisher	Mm00443081_m1
Glucagon TaqMan AOD	Thermo Fisher	Mm00801712_m1
MafB TaqMan AOD	Thermo Fisher	Mm00627481_s1
Arx TaqMan AOD	Thermo Fisher	Mm00545903_m1
NeuroD1 TaqMan AOD	Thermo Fisher	Mm01946604_s1
Foxa2 TaqMan AOD	Thermo Fisher	Mm01976556_s1
Recombinant DNA		
Paupar BAC clone	BAC-PAC resources	RP23-465J7
H2B:GFP	Kanda et al., 1998	Addgene Cat#11680
pL451	Nam and Benezra, 2009	Addgene Cat#22687

pMCS-DTA	Generous gift from Kosuke Yusa, Osaka University, Japan	N/A
Software and Algorithms		
DESeq2	Love et al., 2014	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
GraphPad Prism 7	GraphPad Software	https://www.graphpad.com/scientific-software/prism/
ImageJ	NIH	https://imagej.nih.gov/ij/
Photoshop CC 2018	Adobe	N/A
Illustrator CC 2018	Adobe	N/A
R Software Package 3.3.1	The R Foundation	https://www.r-project.org/
HISAT2 (v2.1.0)	Kim et al., 2013	https://ccb.jhu.edu/software/hisat2/index.shtml
Samtools (v1.4)	Li et al., 2009	http://samtools.sourceforge.net/
HTSeq (v0.10.0)	Anders et al., 2015	https://htseq.readthedocs.io/en/master/install.html
Genomic Regions Enrichment of Annotations Tool (GREAT) (v3.0.0)	McLean et al., 2010	http://great.stanford.edu/public/html/
TopHat2 (v2.1.1)	Kim et al., 2013	http://ccb.jhu.edu/software/tophat/index.shtml
Bowtie2 (v2.2.8)	Langmead and Salzberg, 2012	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
Bedtools (v2.17.0)	Quinlan and Hall, 2010	http://bedtools.readthedocs.io/en/latest/
Heatmapper	Babicki et al., 2016	http://heatmapper.ca/
Coding-Potential Assessment Tool (CPAT) (v1.2.4)	Wang et al., 2013	http://rna-cpat.sourceforge.net/
PhyloCSF	Lin, Jungreis, and Kellis, 2011	https://github.com/mlin/PhyloCSF/wiki
UCSC Genome Browser	Kuhn, Haussler, and Kent, 2013	https://genome.ucsc.edu/
Coding Potential Calculator (CPC)	Kong et al., 2007	http://cpc.cbi.pku.edu.cn/
Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) (v10.0)	Szklarczyk et al., 2017	https://string-db.org/
RBPmap (v1.1)	Paz et al., 2014	http://rbpmap.technion.ac.il/

Replicate Multivariate Analysis of Transcript Splicing (rMATS) (v4.0.2)	Shen et al., 2014	http://rnaseq-mats.sourceforge.net/
IGV	The Broad Institute	http://software.broadinstitute.org/software/igv/
ApE	M. Wayne Davis	http://jorgensen.biology.utah.edu/wayned/ape/

CHAPTER 3: Conclusions and Future Perspectives

Understanding the mechanisms underlying pancreatic islet cell development and function has important implications for the discovery of new therapies for diabetes. For my thesis work, I conducted comparative transcriptome analyses between embryonic mouse pancreas and adult mouse islets and identified several pancreatic lncRNAs that lie in close proximity to essential pancreatic transcription factors, including the Pax6-associated lncRNA *Paupar*. I demonstrated that *Paupar* is enriched in glucagon-producing alpha cells where it promotes the alternative splicing of *Pax6* to an isoform required for activation of essential alpha cell genes. Consistently, deletion of *Paupar* in mice caused dysregulation of PAX6 alpha cell target genes and corresponding alpha cell dysfunction, including blunted glucagon secretion. Taken together, my findings illustrate a distinct mechanism by which a pancreatic lncRNA can coordinate glucose homeostasis by cell-specific regulation of a broadly expressed transcription factor. In this chapter, I will discuss the implications of these findings, caveats of our experimental method, and research questions that remain.

The lncRNA Paupar: illustrating challenges and opportunities

Paupar was first identified in neuroblastoma N2A cells as a single exon lncRNA that regulated genes independently and through direct interaction with PAX6 (Vance et al., 2014). In fact, part of the motivation for focusing on *Paupar* and not other pancreatic lncRNA candidates was that several molecular characteristics of *Paupar* had already been experimentally determined. Interestingly, even though PAX6 has a well-established role in the pancreas, the authors of the original *Paupar* study did not include any pancreatic tissues in their expression panel. This

highlights one of the many challenges in studying lncRNAs in the context of pancreatic islets: (1) isolation of high quality RNA from the pancreas is technically difficult given that the pancreas is one of the richest sources of RNases; (2) pancreatic tissues are therefore rarely included in expression panels used in individual studies or in publically available databases, such as BioGPS and BodyMap; (3) even if the pancreas is in included in a tissue panel, it is often as whole pancreas; (4) islets of Langerhans constitute only ~1-2% of pancreas volume; and (5) lncRNAs are expressed at low levels (an average log10 fold lower than protein coding genes). Taken together, these issues likely contributed to the absence of pancreatic tissues in the Vance study and many other existing studies documenting the tissue specificity of lncRNAs.

When we conducted our own expression analyses, we found that *Paupar* was most highly enriched in pancreatic islets compared to all other tissues examined, including the eye, brain, and N2A cells. Surprisingly, unbiased analysis of the N2A transcriptome showed almost no reads mapping to the *Paupar* locus. In contrast, in the adult islet transcriptome there was robust signal at the Paupar locus. Furthermore, unlike what was reported in N2A cells, in adult islets *Paupar* has three exons and two introns. We also discovered that within alpha cells *Paupar* does not directly interact with PAX6, nor does it regulate *Pax6* expression at the transcript level. These discrepancies between our findings and those published by Vance and colleagues possibly reflect tissue-specific regulation and function. Importantly, these inconsistencies highlight both the importance of identifying the right cell type in which to analyze lncRNA function and the benefit of rigorous and unbiased techniques to investigate lncRNA expression and function.

Noncoding RNAs

While my thesis work focused on lncRNAs, other types of ncRNAs paved the way for our understanding of RNA-mediated gene regulation. The last decade of the 20th century witnessed the birth of the ncRNA revolution, as the first indications of a broad layer of gene regulation ncRNAs began to take shape. In fact, ncRNAs had already been implicated in gene expression; in the 1950s the central role of ribosomal RNAs (rRNAs) and transfer RNAs (tRNAs) in protein synthesis was firmly established. However, it was not until the early 1980s that the first ncRNAs, the small nuclear RNAs (snRNAs), emerged as possible participants in the excision of introns. After their acceptance as building blocks of the spliceosome, the revolution gained huge momentum in the 1990's, when puzzling observations in plants, worms, and flies culminated in the revelation that exogenously supplied ncRNAs could silence the expression of just about any gene of interest through a process aptly named RNA interference (RNAi). Research on these small ncRNAs, now called microRNAs (miRNAs), has shown that these endogenous ~23 nt RNAs play important gene-regulatory roles by pairing to mRNAs to direct their posttranscriptional repression (Bartel 2009). The comparison between miRNAs and lncRNAs activities can tell us much about their respective functions. MicroRNAs are highly conserved across species, likely because their sequence directly determines their downstream mRNA target. In contrast, most lncRNAs are poorly conserved across species, at least at the nucleotide level, likely because they only require short sequences of conservation to maintain secondary structures and interact with target factors. In terms of length, as their name suggests, lncRNAs are much longer than miRNAs, allowing them to have multiple binding motifs for DNA, RNA, and protein. Furthermore, the evolutionary purpose of miRNAs compared to lncRNAs needs to be considered. The metazoan miRNA pathway appears to have been present in the last common ancestor of eukaryotes and continues to defend against viruses and transposons in many

eukaryotes (Bartel 2009). Given the well-established role of miRNAs in developmental biology, perhaps it was important for organisms to evolve a method to reliably inhibit gene expression during cell specification. There are also many more identified lncRNA genes than there are miRNAs, and many different modes of molecular functions, which is consistent with lncRNAs acting through much more diverse lncRNA regulatory mechanisms required to maintain species complexity.

Noncoding RNAs and Species Complexity

Predictions of the estimated number of protein-coding genes in the human genome prior to Human Genome Project ranged from as low as 50,000 to as high as 140,000. To everyone's surprise, sequencing analyses in humans identified only ~20,000 protein-coding genes, similar to other vertebrates such as the mouse and chicken. Furthermore, the nematode worm *C. elegans*, which comprises only 1,000 cells, has 50% more annotated protein-coding genes (~19,300 genes) than far more complex insects (~13,500 genes) and nearly as many genes as currently estimated for vertebrates. Moreover, despite their considerable developmental and neurological complexity, mammals do not appear to have any more protein-coding genes than plants such as *A. thaliana* (~26,000) and rice (~37,000), or protists such as *P. tetraurelia* (~40,000) and *T. thermophila* (~27,000). This post-genomic realization that increased developmental complexity is not reflected in an increased number of protein-coding genes has been termed the G-value paradox (Hahn and Wray, 2002). Part of this paradox can be explained by an increased utilization of alternative splicing, as higher organisms utilize alternative splicing mechanisms more extensively and in a more complicated manner than lower organisms (Nagasaki et al., 2005). Indeed, there is ample evidence that complex organisms utilize a wide range of gene

regulatory processes in addition to alternative splicing, including chromatin architecture, promoter selection, RNA modification and editing, RNA localization, translation, and RNA stability (Taft et al., 2007). However, as deep sequencing technologies have been applied to a wide-range of organisms, it has become clear that the G-value paradox in complex organisms is mostly due to the expansion of cis-acting regulatory elements acting at multiple levels (chromatin architecture, transcription, splicing, mRNA translation, and RNA stability), and the expansion of non-protein-coding genes specifying ncRNAs that fulfill a wide range of regulatory functions (Taft et al., 2007). This concept has interesting relevance for my thesis work on *Paupar* and PAX6. Specifically, Pax6 is highly conserved across bilaterian species, with an established regulatory role in *Drosophila* (eyeless), *C. elegans* (Pax-6), *Xenopus* (pax6), mice (Pax6), and humans (PAX6). In contrast, the lncRNA *Paupar* appears to be conserved only in placental mammals. Taken together with my findings, the presence of a ncRNA (*Paupar*) in more complex species enables PAX6 to achieve more specialized and diverse functions than required in less complex organisms.

LncRNAs, transcription factors, and alternative splicing

Francis Crick's central dogma of molecular biology inspired decades of research structured around the idea that RNA served merely as a messenger between DNA and protein (Crick, 1970). As a result, most cellular processes were studied in a protein-centric manner. This included the regulation of gene expression, which can be divided into two major categories: transcription and translation. Both of these processes tightly control how much of each gene gets expressed in a specific cell type at a specific point in time. Prior to the discovery of regulatory noncoding RNAs, research focused on how proteins orchestrate proper transcription and

translation. Remarkably, in the past decade, lncRNAs have emerged as essential regulators of basically every level of gene expression. For my thesis work, this posed a challenge in that mechanistically, there was no guidelines dictating which pathways might be involved. Conversely, the multitude of possible lncRNA functions made this project incredibly exciting at every turn. This was especially true for our discovery that *Paupar* was involved in alternative splicing.

In the pre-genomics era, alternative splicing (AS) of a gene was considered an unusual event. However, next generation RNA-seq identified that ~95% of multi-exon human genes are alternatively spliced (Pan et al., 2008; Wang et al., 2008), suggesting AS is a key mechanism that may underlie the diversification of proteins encoded in the mammalian genome. Such diversification may be essential for biologic complexity, because the number of protein coding genes is lower than predicted before the genome sequence was known (Kim et al., 2014; Lander et al., 2001). Recently, there has been increasing interest in the role of AS in human diseases, including the dysfunction of islet endocrine cells in diabetes (Juan-Mateu et al., 2015). There are also several studies implicating lncRNAs in different aspects of the AS pathway; however, these studies mainly focused on the role of lncRNAs in forming nuclear structures important for AS (nuclear speckles and paraspeckles), instead of identifying downstream alternatively spliced genes (Tripathi et al., 2010; West et al., 2014). The establishment of lncRNAs in AS is especially interesting from an evolutionary perspective for several reasons; lncRNAs are very cell type specific, allowing for different AS patterns in different cell types; lncRNAs bind easily to other RNA molecules; and lncRNAs are more abundant in complex species and can therefore handle the increased need for highly specialized splicing machinery.

Following our identification of several SRSF proteins as potential *Paupar* binding partners (Chapter 2), we used an unbiased strategy to identify possible downstream targets of *Paupar*-mediated alternative splicing. I initially analyzed the RNA-seq dataset from *Paupar* WT versus *Paupar* KO islets with the goal of trying to find differentially spliced genes that could explain the alpha cell defect seen in *Paupar* KO mice (data not shown). While this analysis yielded a large number of significantly alternatively spliced genes, none stood out genes capable of explaining the *Paupar* KO phenotype. We then took a step back and hypothesized that if *Paupar* was indeed involved in alternative splicing, it was likely controlling the splicing of a gene required for alpha cell development and function. Since my previous splicing analysis was done on whole islets, it was possible that the non-alpha islet cell types were masking changes in alpha cell genes. To circumvent this technical issue, I analyzed published RNA-seq datasets from FACS purified mouse alpha and beta cells (DiGruccio et al., 2016). Surprisingly, *Pax6* emerged as one of the most highly alternatively spliced genes in alpha versus beta cells. This was especially exciting not only because of *Paupar's* genomic proximity to *Pax6*, but also because PAX6 is a classic example of a pancreatic transcription factor that has unique regulatory functions in different islet cell types. For example, several studies have shown that PAX6 has distinct regulatory targets in alpha cells compared to beta cells (Gosmain et al., 2007; Gosmain et al., 2007). We therefore initially hypothesized that the *Pax6 5a* isoform was the "alpha cell" isoform and that the canonical *Pax6* isoform was the "beta cell" isoform. Yet, expression analyses in alphaTC cells showed the presence of both *Pax6* isoforms, which suggested that each isoform had unique intracellular functions, rather than unique intercellular functions. These findings also suggested that perhaps the relative ratio of each isoform is critical for proper gene regulation. This fit well with recent work from Yuval Dor's lab, which demonstrated that PAX6

acts as both a repressor and an activator within pancreatic beta cells (Swisa et al., 2017): PAX6 activated canonical beta cell genes, such as *Ins2, MafA,* and *Nkx6-1,* and repressed beta cell disallowed genes, such as *Gcg* and *Ghrl.*

There were two barriers to teasing out of the function of each *Pax6* isoform. First, we could only quantify the expression of the *Pax6 5a* isoform, which produced a single qRT-PCR product. In contrast, qRT-PCR of the canonical *Pax6* isoform amplified both isoforms. Second, commercially available PAX6 antibodies recognize both isoforms, meaning that we had to infer PAX6 5a activation of *Gcg* and *MafB* based on reduced expression levels. To summarize our findings from Chapter 2, we next showed that in alpha cells, *Paupar* promotes the alternative splicing of *Pax6* pre-mRNA to the *Pax6 5a* isoform, which in turn is required for the activation of essential alpha cell genes. These findings have several implications for our understanding of the mechanisms underlying regulation of gene expression. First, methods used to understand lncRNA function should be as unbiased as possible. LncRNAs can physically interact with DNA, other RNAs, and protein, and although these sequencing and proteomic experiments can be costly and time-consuming, visualizing the lncRNA interactome will provide the lncRNA field with the information it needs to build an evidence-based classification system. Second, differential expression analysis does not always tell the whole story behind a complex phenotype. Prior to our discovery that *Paupar* interacted with splicing factors, we had ruled out that *Paupar* regulated *Pax6* because we did not see misregulation of *Pax6* in *Paupar*-deficient cells at the whole gene level. This study taught me how incredibly complex biology can be and how important it is to allow the data to inform the hypothesis and not the other way around. Finally, this study adds to the foundation of knowledge showing that lncRNAs are key regulators

of gene expression, and provides early evidence that lncRNAs are essential for islet development and function. Furthermore, I hope that these findings motivate all scientists, even the skeptics, to reexamine their biological questions in the context of a role for RNA.

Future directions

My thesis furthered our understanding of the mechanisms mediating alpha cell development and function; however, several important research questions remain. First, while *Paupar* confers the alpha cell specific function on *Pax6*, what is the mechanism underlying the cell specific expression of *Paupar*? In other words, what regulates *Paupar*? The onset of *Paupar* expression during alpha cell maturation suggests that *Paupar* activation is mediated by a maturation factor. Interestingly, establishment and maintenance of alpha and beta cell identity is regulated by a complex interplay of transcription factors (Benner et al., 2014). Comparative transcriptome analyses of postnatal maturing alpha and beta cells showed that immature alpha cells are first marked by *Arx, Isl1, NeuroD1*, and *Pax6* (Benner et al., 2014). A few days later, maturation induces the expression of *Gcg, Irx1, Irx2, MafB*, and *Brn4* (Benner et al., 2014). Our findings that *Paupar* activates the expression of *Gcg* and *MafB* suggests that activation of *Paupar* occurs in early immature alpha cells, possibly by the transcription factors *Arx, Isl1, NeuroD1*, or *Pax6*. Future experiments could assess *Paupar* expression following knockdown of these transcription factors in alphaTCs. We could also use bioinformatics to look for transcription factor binding sites in the putative *Paupar* promoter. Unfortunately, information that may elucidate alpha cell specific promoters and enhancers, such as ChIP-seq for histone marks, is currently unavailable for islet alpha cells. Hopefully this will change soon with the increased attention on islet alpha cells in diabetes and improved sequencing technologies better suited for rare cell populations.

A major goal of our research is to better understand and treat diabetes in humans; however, how well our findings in mice translate to humans is not always clear. One of the characteristics of *Paupar* that suggested it warranted further study was its high level of conservation in humans: *Paupar* is conserved both syntentically and by a majority (84%) of its nucleotide sequence. I have confirmed that *PAUPAR* is expressed in human islets, but it seems to exhibit slightly higher expression in beta cells compared to alpha cells (data not shown). This finding is interesting given that unlike in mice, MAFB is expressed in both human alpha and beta cells. Perhaps the mechanism of *Paupar*-mediated activation of *MafB* is conserved in humans and that is why *Paupar* is also expressed in beta cells. Future experiments could test this by knocking down *Paupar* in the EndoB1C human beta cell line or during *in vitro* hESC to beta cell differentiation. Experiments using the latter would also allow us to test whether human *PAUPAR* was required for the specification of human alpha and/or beta cells.

LncRNAs have been shown to be attractive candidates for drug targeting, which could be exploited for developing new diabetes therapies. For example, lncRNAs can be targeted by antisense oligonucleotides (ASOs), which is technically simpler than small molecule screening or inhibitory antibody development. Furthermore, the highly tissue specific expression of lncRNAs also allows the use of lower doses of targeting molecules, thus alleviating drug toxicity. In the context of *Paupar*, due to the recently renewed appreciation of the role of hyperglucagonemia in diabetes (Brissova et al., 2018), there is a great interest in diabetes therapies that can reduce glucagon secretion. Significantly, our finding that knockdown of *Paupar* causes reduced glucagon secretion in vivo suggests reducing *Paupar* in diabetic islets could be a promising method to reduce diabetic hyperglycemia. As previously mentioned, we

would first have to show that *Paupar*-mediated regulation of glucagon secretion is conserved in humans. We could then knockdown *Paupar* in human diabetic islets and test whether this corrected hyperglucagonemia. Taken together, *Paupar* is an appealing gene candidate for specific treatment of hyperglycemia seen in diabetes.

Concluding Remarks

Noncoding RNAs are increasingly recognized as essential regulators of a wide range of biological processes. LncRNAs are the most abundant type of ncRNA and a large number of islet-specific lncRNAs have been identified; however, their functional characterization in the pancreas is lacking. During my thesis work, I conducted extensive molecular and functional characterization of an alpha cell enriched lncRNA. My findings suggest that lncRNAs could represent important cell restricted therapeutic targets to regulate cell-specific functions of more widely expressed proteins. Expansion of these findings to additional lncRNAs will further our understanding of islet development and function and help identify novel therapeutic targets for the treatment of diabetes.

REFERENCES

Ackermann, A.M., Wang, Z., Schug, J., Naji, A., Kaestner, K.H. (2016). Integration of ATAC-seq and RNA-seq identifies human alpha cell and beta cell signature genes. Molecular Metabolism *5,* 233-244.

Adams, B.D., Parsons, C., Walker, L., Zhang, W.C., and Slack, F.J. (2017). Targeting noncoding RNAs in disease. J Clin Invest 127, 761-771.

Ahlgren, U., Jonsson, J., and Edlund, H. (1996). The morphogenesis of the pancreatic mesenchyme is uncoupled from that of the pancreatic epithelium in IPF1/PDX1-deficient mice. Development 122, 1409-1416.

Ahmad, Z., Rafeeq, M., Collombat, P., and Mansouri, A. (2015). Pax6 Inactivation in the Adult Pancreas Reveals Ghrelin as Endocrine Cell Maturation Marker. PLoS One 10, e0144597.

Ahnfelt-Ronne, J., Jorgensen, M.C., Klinck, R., Jensen, J.N., Fuchtbauer, E.M., Deering, T., MacDonald, R.J., Wright, C.V., Madsen, O.D., and Serup, P. (2012). Ptf1a-mediated control of Dll1 reveals an alternative to the lateral inhibition mechanism. Development 139, 33-45.

Akerman, I., Tu, Z., Beucher, A., Rolando, D.M.Y., Sauty-Colace, C., Benazra, M., Nakic, N., Yang, J., Wang, H., Pasquali, L., et al. (2017). Human Pancreatic beta Cell lncRNAs Control Cell-Specific Regulatory Networks. Cell Metab 25, 400-411.

Allen, F. (1913). Studies Concerning Glycosuria and Diabetes. WM Leonard Boston.

Alvarez, M.L., and DiStefano, J.K. (2011). Functional characterization of the plasmacytoma variant translocation 1 gene (PVT1) in diabetic nephropathy. PLoS One 6, e18671.

Amaral, P.P., Clark, M.B., Gascoigne, D.K., Dinger, M.E., and Mattick, J.S. (2011). lncRNAdb: a reference database for long noncoding RNAs. Nucleic Acids Research 39, D146-D151.

American Diabetes Association. (2018). Type 1. http://www.diabetes.org/diabetes-basics/type-1.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166-169.

Anderson, K.R., Torres, C.A., Solomon, K., Becker, T.C., Newgard, C.B., Wright, C.V., Hagman, J., and Sussel, L. (2009). Cooperative transcriptional regulation of the essential pancreatic islet gene NeuroD1 (beta2) by Nkx2.2 and neurogenin 3. J Biol Chem 284, 31236-31248.

Andersson, S.A., Olsson, A.H., Esguerra, J.L., Heimann, E., Ladenvall, C., Edlund, A., Salehi, A., Taneera, J., Degerman, E., Groop, L., et al. (2012). Reduced insulin secretion correlates with decreased expression of exocytotic genes in pancreatic islets from patients with type 2 diabetes. Mol Cell Endocrinol 364, 36-45.

Annes, J.P., Ryu, J.H., Lam, K., Carolan, P.J., Utz, K., Hollister-Lock, J., Arvanites, A.C., Rubin, L.L., Weir, G., and Melton, D.A. (2012). Adenosine kinase inhibition selectively promotes rodent and porcine islet beta-cell replication. Proc Natl Acad Sci U S A 109, 3915-3920.

Apelqvist, A., Li, H., Sommer, L., Beatus, P., Anderson, D.J., Honjo, T., de Angelis, M.H., Lendahl, U., and Edlund, H. (1999). Notch signalling controls pancreatic cell differentiation. Nature 400, 877-881.

Arda, H.E., Benitez, C.M., and Kim, S.K. (2013). Gene regulatory networks governing pancreas development. Dev Cell 25, 5-13.

Arieff AJ, Crawford J, Adams J, Smith D. Glucagon in insulin coma therapy: its use in a small psychiatric unit of a general hospital. Q Bull Northwest Univ Med Sch 34: 7–10, 1960.

Arnes, L., Akerman, I., Balderes, D.A., Ferrer, J., and Sussel, L. (2016). βlinc1 encodes a long noncoding RNA that regulates islet β-cell formation and function. Genes and Development 30, 502-507.

Arnes, L., Liu, Z., Wang, J., Carlo Maurer, H., Sagalovskiy, I., Sanchez-Martin, M., Bommakanti, N., Garofalo, D.C., Balderes, D.A., Sussel, L., et al. (2018). Comprehensive characterisation of compartment-specific long non-coding RNAs associated with pancreatic ductal adenocarcinoma. Gut. 0, 1-16.

Artner, I., Blanchi, B., Raum, J.C., Guo, M., Kaneko, T., Cordes, S., Sieweke, M., and Stein, R. (2007). MafB is required for islet beta cell maturation. Proc Natl Acad Sci U S A 104, 3853-3858.

Artner, I., Lay, J.L., Hang, Y., Elghazi, L., Schisler, J.C., Henderson, E., Sosa-Pineda, B., and Stein, R. (2006). An Activator of the Glucagon Gene Expressed in Developing Islet Diabetes 55, 297-304.

Ashcroft, F.M., Harrison, D.E., and Ashcroft, S.J.H. (1984). Glucose induces closure of single potassium channels in isolated rat pancreatic beta-cells. Nature 312, 446-448.

Ashery-Padan, R., Zhou, X., Marquardt, T., Herrera, P., Toube, L., Berry, A., and Gruss, P. (2004). Conditional inactivation of Pax6 in the pancreas causes early onset of diabetes. Dev Biol 269, 479-488.

Atkinson, M.A., Eisenbarth, G.S., and Michels, A.W. (2014). Type 1 diabetes. The Lancet 383, 69-82.

Awata, T., Yamashita, H., Kurihara, S., Morita-Ohkubo, T., Miyashita, Y., Katayama, S., Mori, K., Yoneya, S., Kohda, M., Okazaki, Y., et al. (2014). A genome-wide association study for diabetic retinopathy in a Japanese population: potential association with a long intergenic non-coding RNA. PLoS One 9, e111715.

Babicki, S., Arndt, D., Marcu, A., Liang, Y., Grant, J.R., Maciejewski, A., and Wishart, D.S. (2016). Heatmapper: web-enabled heat mapping for all. Nucleic Acids Res 44, W147-153.

Baeyens, L., Lemper, M., Leuckx, G., De Groef, S., Bonfanti, P., Stange, G., Shemer, R., Nord, C., Scheel, D.W., Pan, F.C., et al. (2014). Transient cytokine treatment induces acinar cell reprogramming and regenerates functional beta cell mass in diabetic mice. Nat Biotechnol 32, 76-83.

Banting, F. and Best, C. (1922). The internal secretion of the pancreas. J. Lab. Clin. Med. *7*, 251–266.

Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. Cell 136, 215-233.

Barton, F.B., Rickels, M.R., Alejandro, R., Hering, B.J., Wease, S., Naziruddin, B., Oberholzer, J., Odorico, J.S., Garfinkel, M.R., Levy, M., et al. (2012). Improvement in outcomes of clinical islet transplantation: 1999-2010. Diabetes Care 35, 1436-1445.

Baum, J., Simons B.E., Unger R.H., Madison L.L. (1962). Localization of glucagon in the alpha cells in the pancreatic islet by immunofluorescent technics. Diabetes. *11*, 371-374.

Bassett, A.R., Azzam, G., Wheatley, L., Tibbit, C., Rajakumar, T., McGowan, S., Stanger, N., Ewels, P.A., Taylor, S., Ponting, C.P., et al. (2014). Understanding functional miRNA-target interactions in vivo by site-specific genome engineering. Nat Commun 5, 4640.

Batista, P.J., and Chang, H.Y. (2013). Long noncoding RNAs: cellular address codes in development and disease. Cell 152, 1298-1307.

Bellin, M.D., Barton, F.B., Heitman, A., Harmon, J.V., Kandaswamy, R., Balamurugan, A.N., Sutherland, D.E., Alejandro, R., and Hering, B.J. (2012). Potent induction immunotherapy promotes long-term insulin independence after islet transplantation in type 1 diabetes. Am J Transplant 12, 1576-1583.

Ben-Othman, N., Vieira, A., Courtney, M., Record, F., Gjernes, E., Avolio, F., Hadzic, B., Druelle, N., Napolitano, T., Navarro-Sanz, S., et al. (2017). Long-Term GABA Administration Induces Alpha Cell-Mediated Beta-like Cell Neogenesis. Cell 168, 73-85 e11.

Benner, C., van der Meulen, T., Caceres, E., Tigyi, K., Donaldson, C.J., and Huising, M.O. (2014). The transcriptional landscape of mouse beta cells compared to human beta cells reveals notable species differences in long non-coding RNA and protein-coding gene expression. BMC Genomics 15, 1-16.

Bernstein, E., Kim, S.Y., Carmell, M.A., Murchison, E.P., Alcorn, H., Li, M.Z., Mills, A.A., Elledge, S.J., Anderson, K.V., and Hannon, G.J. (2003). Dicer is essential for mouse development. Nat Genet 35, 215-217.

Bliss, M. (1982). Banting's, Best's, and Collip's Accounts of the Discovery of Insulin. Bulletin of the History of Medicine 56, 554-568.

Bliss, M. (1982). The Discovery of Insulin. The University of Chicago Press 238.

Blodgett, D.M., Nowosielska, A., Afik, S., Pechhold, S., Cura, A.J., Kennedy, N.J., Kim, S., Kucukural, A., Davis, R.J., Kent, S.C., et al. (2015). Novel Observations From Next-Generation RNA Sequencing of Highly Purified Human Adult and Fetal Islet Cell Subsets. Diabetes 64, 3172-3181.

Bonner-Weir, S., Sullivan, B.A., and Weir, G.C. (2015). Human Islet Morphology Revisited: Human and Rodent Islets Are Not So Different After All. J Histochem Cytochem 63, 604-612.

Bouwens, L., Houbracken, I., and Mfopou, J.K. (2013). The use of stem cells for pancreatic regeneration in diabetes mellitus. Nat Rev Endocrinol 9, 598-606.

Bramswig, N.C., Everett, L.J., Schug, J., Dorrell, C., Liu, C., Luo, Y., Streeter, P.R., Naji, A., Grompe, M., and Kaestner, K.H. (2013). Epigenomic plasticity enables human pancreatic alpha to beta cell reprogramming. J Clin Invest 123, 1275-1284.

Briant, L.J.B., Reinbothe, T.M., Spiliotis, I., Miranda, C., Rodriguez, B., and Rorsman, P. (2018). delta-cells and beta-cells are electrically coupled and regulate alpha-cell activity via somatostatin. J Physiol 596, 197-215.

Brissova, M., Fowler, M.J., Nicholson, W.E., Chu, A., Hirshberg, B., Harlan, D.M., and Powers, A.C. (2005). Assessment of human pancreatic islet architecture and composition by laser scanning confocal microscopy. J Histochem Cytochem 53, 1087-1097.

Brissova, M., Haliyur, R., Saunders, D., Shrestha, S., Dai, C., Blodgett, D.M., Bottino, R., Campbell-Thompson, M., Aramandla, R., Poffenberger, G., et al. (2018). alpha Cell Function and Gene Expression Are Compromised in Type 1 Diabetes. Cell Rep 22, 2667-2676.

Bromer WW, Sinn LG, Staub A, Behrens OK. The amino acid sequence of glucagon. Diabetes 6: 234 –238, 1957.

Bürger, M., Brandt, W. (1935). Uber das Glukagon (die hyperglykämisierende Substanz der Pankreas). Z Ges Exp Med *96*, 375.

Bürger, M., Kramer, H. (1929). Primäre Hyperglykämie und Glykogenveränderung der Leber als Folge intraportaler Insulininjektion nach Untersuchungen am Hund. Z Ges Exp Med *67*: 441.

Burlison, J.S., Long, Q., Fujitani, Y., Wright, C.V., and Magnuson, M.A. (2008). Pdx-1 and Ptf1a concurrently determine fate specification of pancreatic multipotent progenitor cells. Dev Biol 316, 74-86.

Cabili, M.N., Dunagin, M.C., McClanahan, P.D., Biaesch, A., Padovan-Merhar, O., Regev, A., Rinn, J.L., and Raj, A. (2015). Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. Genome Biol 16, 20.

Cabrera, O., Berman, D.M., Kenyon, N.S., Ricordi, C., Berggren, P.O., and Caicedo, A. (2006). The unique cytoarchitecture of human pancreatic islets has implications for islet cell function. Proc Natl Acad Sci U S A 103, 2334-2339.

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., et al. (2005). The transcriptional landscape of the mammalian genome. Science 309, 1559-1563.

Carrero, J.A., Calderon, B., Towfic, F., Artyomov, M.N., and Unanue, E.R. (2013). Defining the transcriptional and cellular landscape of type 1 diabetes in the NOD mouse. PLoS One 8, e59701.

Carter, G., Miladinovic, B., Patel, A.A., Deland, L., Mastorides, S., and Patel, N.A. (2015). Circulating long noncoding RNA GAS5 levels are correlated to prevalence of type 2 diabetes mellitus. BBA Clin 4, 102-107.

Cavelti-Weder, C., Li, W., Zumsteg, A., Stemann, M., Yamada, T., Bonner-Weir, S., Weir, G., and Zhou, Q. (2015). Direct Reprogramming for Pancreatic Beta-Cells Using Key Developmental Genes. Curr Pathobiol Rep 3, 57-65.

Centers for Disease Control and Prevention. National Diabetes Statistics Report, (2017) Atlanta, GA

Ceranowicz, P., Cieszkowski, J., Warzecha, Z., Kusnierz-Cabala, B., and Dembinski, A. (2015). The Beginnings of Pancreatology as a Field of Experimental and Clinical Medicine. Biomed Res Int 2015, 128095.

Cerasi, E., Efendic, S., and Luft, R. (1973). Dose-response relation between plasma-insulin and blood-glucose levels during oral glucose loads in prediabetic and diabetic subjects. The Lancet 1, 794-796.

Cerda-Esteban, N., Naumann, H., Ruzittu, S., Mah, N., Pongrac, I.M., Cozzitorto, C., Hommel, A., Andrade-Navarro, M.A., Bonifacio, E., and Spagnoli, F.M. (2017). Stepwise reprogramming of liver cells to a pancreas progenitor state by the transcriptional regulator Tgif2. Nat Commun 8, 14127.

Chakravarthy, H., Gu, X., Enge, M., Dai, X., Wang, Y., Damond, N., Downie, C., Liu, K., Wang, J., Xing, Y., et al. (2017). Converting Adult Pancreatic Islet alpha Cells into beta Cells by Targeting Both Dnmt1 and Arx. Cell Metab 25, 622-634.

Chauhan, B.K., Yang, Y., Cveklova, K., and Cvekl, A. (2004). Functional Properties of Natural Human PAX6 and PAX6(5a) Mutants. Invest Ophthalmol Vis Sci 45, 385-392.

Chen, Y.J., Finkbeiner, S.R., Weinblatt, D., Emmett, M.J., Tameire, F., Yousefi, M., Yang, C., Maehr, R., Zhou, Q., Shemer, R., et al. (2014). De novo formation of insulin-producing "neo-beta cell islets" from intestinal crypts. Cell Rep 6, 1046-1058.

Chera, S., Baronnier, D., Ghila, L., Cigliola, V., Jensen, J.N., Gu, G., Furuyama, K., Thorel, F., Gribble, F.M., Reimann, F., et al. (2014). Diabetes recovery by age-dependent conversion of pancreatic delta-cells into insulin producers. Nature 514, 503-507.

Chu, C., Qu, K., Zhong, F.L., Artandi, S.E., and Chang, H.Y. (2011). Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. Mol Cell 44, 667-678.

Chu, C., Zhang, Q.C., da Rocha, S.T., Flynn, R.A., Bharadwaj, M., Calabrese, J.M., Magnuson, T., Heard, E., and Chang, H.Y. (2015). Systematic discovery of Xist RNA binding proteins. Cell 161, 404-416.

Churchill, A.J., Gutierrez, G.D., Singer, R.A., Lorberbaum, D.S., Fischer, K.A., and Sussel, L. (2017). Genetic evidence that Nkx2.2 acts primarily downstream of Neurog3 in pancreatic endocrine lineage development. Elife 6.

Collombat, P., Mansouri, A., Hecksher-Sorensen, J., Serup, P., Krull, J., Gradwohl, G., and Gruss, P. (2003). Opposing actions of Arx and Pax4 in endocrine pancreas development. Genes Dev 17, 2591-2603.

Collombat, P., Xu, X., Ravassard, P., Sosa-Pineda, B., Dussaud, S., Billestrup, N., Madsen, O.D., Serup, P., Heimberg, H., and Mansouri, A. (2009). The ectopic expression of Pax4 in the mouse pancreas converts progenitor cells into alpha and subsequently beta cells. Cell 138, 449-462.

Conarello, S.L., Jiang, G., Mu, J., Li, Z., Woods, J., Zycband, E., Ronan, J., Liu, F., Roy, R.S., Zhu, L., et al. (2007). Glucagon receptor knockout mice are resistant to diet-induced obesity and streptozotocin-mediated beta cell loss and hyperglycaemia. Diabetologia 50, 142-150.

Cook, D.L., and Hales, N. (1984). Intracellular ATP directly blocks K+ channels in pancreatic B-cells. Nature 311, 271-272.

Courtney, M., Gjernes, E., Druelle, N., Ravaud, C., Vieira, A., Ben-Othman, N., Pfeifer, A., Avolio, F., Leuckx, G., Lacas-Gervais, S., et al. (2013). The inactivation of Arx in pancreatic alpha-cells triggers their neogenesis and conversion into functional beta-like cells. PLoS Genet 9, e1003934.

Crick, F. (1970). Central Dogma of Molecular Biology. Nature *227*, 561-563.

Cropano, C., Santoro, N., Groop, L., Dalla Man, C., Cobelli, C., Galderisi, A., Kursawe, R., Pierpont, B., Goffredo, M., and Caprio, S. (2017). The rs7903146 Variant in the TCF7L2 Gene Increases the Risk of Prediabetes/Type 2 Diabetes in Obese Adolescents by Impairing beta-Cell Function and Hepatic Insulin Sensitivity. Diabetes Care 40, 1082-1089.

Cui, Y., Hu, D., Markillie, L.M., Chrisler, W.B., Gaffrey, M.J., Ansong, C., Sussel, L., and Orr, G. (2018). Fluctuation localization imaging-based fluorescence in situ hybridization (fliFISH) for accurate detection and counting of RNA copies in single cells. Nucleic Acids Res 46, e7.

Curry, D.L., Bennett, L.L., Grodsky, G.M. (1968). Dynamics of Insulin Secretion by the Perfused Rat Pancreas. Endocrinology. *83*, 572-584.

D'Amour, K.A., Bang, A.G., Eliazer, S., Kelly, O.G., Agulnick, A.D., Smart, N.G., Moorman, M.A., Kroon, E., Carpenter, M.K., and Baetge, E.E. (2006). Production of pancreatic hormone–expressing endocrine cells from human embryonic stem cells. Nature Biotechnology 24, 1392-1401.

Dadon, D., Tornovksy-Babaey, S., Furth-Lavi, J., Ben-Zvi, D., Ziv, O., Schyr-Ben-Haroush, R., Stolovich-Rain, M., Hija, A., Porat, S., Granot, Z., et al. (2012). Glucose metabolism: key endogenous regulator of β-cell replication and survival. Diabetes, Obesity and Metabolism 14, 101-108.

Dai, C., Hang, Y., Shostak, A., Poffenberger, G., Hart, N., Prasad, N., Phillips, N., Levy, S.E., Greiner, D.L., Shultz, L.D., et al. (2017). Age-dependent human beta cell proliferation induced by glucagon-like peptide 1 and calcineurin signaling. J Clin Invest 127, 3835-3844.

Darnell, R. (2012). CLIP (cross-linking and immunoprecipitation) identification of RNAs bound by a specific protein. Cold Spring Harb Protoc 2012, 1146-1160.

de Gonzalo-Calvo, D., Kenneweg, F., Bang, C., Toro, R., van der Meer, R.W., Rijzewijk, L.J., Smit, J.W., Lamb, H.J., Llorente-Cortes, V., and Thum, T. (2016). Circulating long-non coding RNAs as biomarkers of left ventricular diastolic function and remodelling in patients with well-controlled type 2 diabetes. Sci Rep 6, 37354.

Dean, E.D., Li, M., Prasad, N., Wisniewski, S.N., Von Deylen, A., Spaeth, J., Maddison, L., Botros, A., Sedgeman, L.R., Bozadjieva, N., et al. (2017). Interrupted Glucagon Signaling Reveals Hepatic alpha Cell Axis and Role for L-Glutamine in alpha Cell Proliferation. Cell Metab 25, 1362-1373 e1365.

DeFronzo, R.A., Ferrannini, E., Groop, L., Henry, R.R., Herman, W.H., Holst, J.J., Hu, F.B., Kahn, C.R., Raz, I., Shulman, G.I., et al. (2015). Type 2 diabetes mellitus. Nat Rev Dis Primers 1, 15019.

DeFronzo, R.A., Hompesch, M., Kasichayanula, S., Liu, X., Hong, Y., Pfister, M., Morrow, L.A., Leslie, B.R., Boulton, D.W., Ching, A., et al. (2013). Characterization of Renal Glucose Reabsorption in Response to Dapagliflozin in Healthy Subjects and Subjects With Type 2 Diabetes. Emerging Technologies and Therapeutics 36, 3169-3176.

DeFronzo, R.A., and Tripathy, D. (2009). Skeletal muscle insulin resistance is the primary defect in type 2 diabetes. Diabetes Care 32 Suppl 2, S157-163.

DeFronzo, R.A., Triplitt, C., Qu, Y., Lewis, M.S., Maggs, D., and Glass, L.C. (2010). Effects of exenatide plus rosiglitazone on beta-cell function and insulin sensitivity in subjects with type 2 diabetes on metformin. Diabetes Care 33, 951-957.

Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., Guernec, G., Martin, D., Merkel, A., Knowles, D.G., et al. (2012). The GENCODE v7 catalog of human long

noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res 22, 1775-1789.

Dhawan, S., Georgia, S., Tschen, S.I., Fan, G., and Bhushan, A. (2011). Pancreatic beta cell identity is maintained by DNA methylation-mediated repression of Arx. Dev Cell 20, 419-429.

Diaz-Valencia, P.A., Bougneres, P., and Valleron, A.J. (2015). Global epidemiology of type 1 diabetes in young adults and adults: a systematic review. BMC Public Health 15, 255.

DiGruccio, M.R., Mawla, A.M., Donaldson, C.J., Noguchi, G.M., Vaughan, J., Cowing-Zitron, C., van der Meulen, T., and Huising, M.O. (2016). Comprehensive alpha, beta and delta cell transcriptomes reveal that ghrelin selectively activates delta cells and promotes somatostatin release from pancreatic islets. Mol Metab 5, 449-458.

Doerks, T., Copley, R.R., Schultz, J., Ponting, C.P., and Bork, P. (2002). Systematic identification of novel protein domain families associated with nuclear functions. Genome Res 12, 47-56.

Dor, Y., Brown, J., Martinez, O.I., and Melton, D.A. (2004). Adult pancreatic b-cells are formed by self-duplication rather than stem-cell differentiation. Nature 429, 41-46.

Dorrell, C., Schug, J., Lin, C.F., Canaday, P.S., Fox, A.J., Smirnova, O., Bonnah, R., Streeter, P.R., Stoeckert, C.J., Jr., Kaestner, K.H., et al. (2011). Transcriptomes of the major human pancreatic cell types. Diabetologia 54, 2832-2844.

Du, A., Hunter, C.S., Murray, J., Noble, D., Cai, C.L., Evans, S.M., Stein, R., and May, C.L. (2009). Islet-1 is required for the maturation, proliferation, and survival of the endocrine pancreas. Diabetes 58, 2059-2069.

Dunning, B.E., and Gerich, J.E. (2007). The role of alpha-cell dysregulation in fasting and postprandial hyperglycemia in type 2 diabetes and therapeutic implications. Endocr Rev 28, 253-283.

Ediger, B.N., Lim, H.W., Juliana, C., Groff, D.N., Williams, L.T., Dominguez, G., Liu, J.H., Taylor, B.L., Walp, E.R., Kameswaran, V., et al. (2017). LIM domain-binding 1 maintains the terminally differentiated state of pancreatic beta cells. J Clin Invest 127, 215-229.

Efrat, S., and Russ, H.A. (2012). Making beta cells from adult tissues. Trends Endocrinol Metab 23, 278-285.

Eisenbarth, G.S. (1986). Type 1 Diabetes Mellitus, A Chronic Autoimmune Disease. New England Journal of Medicine 314, 1360-1368.

Engreitz, J.M., Pandya-Jones, A., McDonel, P., Shishkin, A., Sirokman, K., Surka, C., Kadri, S., Xing, J., Goren, A., Lander, E.S., et al. (2013). The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. Science 341, 1237973.

Epstein, J.A., Glaser, T., Cai, J., Jepeal, L., Walton, D.S., and Maas, R.L. (1994). Two independent and interactive DNA-binding subdomains of the Pax6 paired domain are regulated by alternative splicing. Genes and Development 8, 2022-2034.

Erlandsen, S.L., Hegre, O.D., Parsons, J.A., McEvoy, R.C., and Elde, R.P. (1976). Pancreatic Isle Cell Hormones Distribution of Cell Types in the Islet and Evidence for the Presence of Somatostatin and Gastrin Within the D Cell. Journal of Histochemistry and Cytochemistry 24, 883-897.

Esguerra, J.L.S., and Eliasson, L. (2014). Functional implications of long non-coding RNAs in the pancreatic islets of Langerhans. Frontiers in Genetics. *5*, 1-9.

Esquibel AJ, Kurland AA, Mendelsohn D. The use of glucagon in terminating insulin coma. Dis Nerv Syst 19: 485– 486, 1958.

Fadista, J., Vikman, P., Laakso, E.O., Mollet, I.G., Esguerra, J.L., Taneera, J., Storm, P., Osmark, P., Ladenvall, C., Prasad, R.B., et al. (2014). Global genomic and transcriptomic analysis of human pancreatic islets reveals novel genes influencing glucose metabolism. Proc Natl Acad Sci U S A 111, 13924-13929.

Flannick, J., Thorleifsson, G., Beer, N.L., Jacobs, S.B., Grarup, N., Burtt, N.P., Mahajan, A., Fuchsberger, C., Atzmon, G., Benediktsson, R., et al. (2014). Loss-of-function mutations in SLC30A8 protect against type 2 diabetes. Nat Genet 46, 357-363.

Fort, K.L., van de Waterbeemd, M., Boll, D., Reinhardt-Szyba, M., Belov, M.E., Sasaki, E., Zschoche, R., Hilvert, D., Makarov, A.A., and Heck, A.J.R. (2017). Expanding the structural analysis capabilities on an Orbitrap-based mass spectrometer for large macromolecular complexes. Analyst 143, 100-105.

Freychet, L., Desplanque, N., Zirinis, P., Rizkalla, S.W., Basdevant, A., Tchobroutsky, G., and Slama, G. (1988). Effect of Intranasal Glucagon On Blood Glucose Levels in Healthy Subjects and Hypoglycaemic Patients With Insulin-Dependent Diabetes. The Lancent 1, 1364-1366.

Friedman, J.M. (2010). A tale of two hormones. Nat Med 16, 1100-1106.

Fuchsberger, C., Flannick, J., Teslovich, T.M., Mahajan, A., Agarwala, V., Gaulton, K.J., Ma, C., Fontanillas, P., Moutsianas, L., McCarthy, D.J., et al. (2016). The genetic architecture of type 2 diabetes. Nature 536, 41-47.

Fujitani, Y., Fujitani, S., Luo, H., Qiu, F., Burlison, J., Long, Q., Kawaguchi, Y., Edlund, H., MacDonald, R.J., Furukawa, T., et al. (2006). Ptf1a determines horizontal and amacrine cell fates during mouse retinal development. Development 133, 4439-4450.

Furuta, M., Zhou, A., Webb, G., Carroll, R., Ravazzola, M., Orci, L., and Steiner, D.F. (2001). Severe defect in proglucagon processing in islet A-cells of prohormone convertase 2 null mice. J Biol Chem 276, 27197-27202.

Gannon, M., Ables, E.T., Crawford, L., Lowe, D., Offield, M.F., Magnuson, M.A., and Wright, C.V. (2008). pdx-1 function is specifically required in embryonic beta cells to generate appropriate numbers of endocrine cell types and maintain glucose homeostasis. Dev Biol 314, 406-417.

Gao, T., McKenna, B., Li, C., Reichert, M., Nguyen, J., Singh, T., Yang, C., Pannikar, A., Doliba, N., Zhang, T., et al. (2014). Pdx1 maintains beta cell identity and function by repressing an alpha cell program. Cell Metab 19, 259-271.

Gauthier, B.R., Gosmain, Y., Mamin, A., and Philippe, J. (2007). The beta-cell specific transcription factor Nkx6.1 inhibits glucagon gene transcription by interfering with Pax6. Biochem J 403, 593-601.

Gelling, R.W., Du, X.Q., Dichmann, D.S., Romer, J., Huang, H., Cui, L., Obici, S., Tang, B., Holst, J.J., Fledelius, C., et al. (2003). Lower blood glucose, hyperglucagonemia, and pancreatic alpha cell hyperplasia in glucagon receptor knockout mice. Proc Natl Acad Sci U S A 100, 1438-1443.

Gepts, W. (1965). Pathologic Anatomy of the Pancreas in Juvenile Diabetes Mellitus. Diabetes 14, 619-633.

Gerich, J.E., Charles, M.A., and Grodsky, G.M. (1974). Characterization of the effects of arginine and glucose on glucagon and insulin release from the perfused rat pancreas. J Clin Invest 54, 833-841.

Gerich, J.E., Meyer, C., Woerle, H.J., and Stumvoll, M. (2001). Renal Gluconeogenesis. Diabetes Care 24, 382-391.

Giroud, M., and Scheideler, M. (2017). Long Non-Coding RNAs in Metabolic Organs and Energy Homeostasis. Int J Mol Sci 18.

Gittes, G.K. (2009). Developmental biology of the pancreas: a comprehensive review. Dev Biol 326, 4-35.

Goke, B. (2008). Islet cell function: alpha and beta cells--partners towards normogylcaemia. International Journal of clinical practice. 62, 2-7.

Golosow, N., and Grobstein, C. (1962). Epitheliomesenchymal Interaction in Pancreatic Morphogenesis. Developmental Biology 4, 242-255.

Gosmain, Y., Avril, I., Mamin, A., and Philippe, J. (2007). Pax-6 and c-Maf functionally interact with the alpha-cell-specific DNA element G1 in vivo to promote glucagon gene expression. J Biol Chem 282, 35024-35034.

Gosmain, Y., Katz, L.S., Masson, M.H., Cheyssac, C., Poisson, C., and Philippe, J. (2012). Pax6 is crucial for beta-cell function, insulin biosynthesis, and glucose-induced insulin secretion. Mol Endocrinol 26, 696-709.

Gosmain, Y., Marthinet, E., Cheyssac, C., Guerardel, A., Mamin, A., Katz, L.S., Bouzakri, K., and Philippe, J. (2010). Pax6 controls the expression of critical genes involved in pancreatic {alpha} cell differentiation and function. J Biol Chem 285, 33381-33393.

Gouzi, M., Kim, Y.H., Katsumoto, K., Johansson, K., and Grapin-Botton, A. (2011). Neurogenin3 initiates stepwise delamination of differentiating endocrine cells during pancreas development. Dev Dyn 240, 589-604.

Gradwohl, G., Dierich, A., LeMeur, M., and Guillemot, F. (2000). neurogenin3 is required for the development of the four endocrine cell lineages of the pancreas. PNAS 97, 1607-1611.

Gregg, B.E., Moore, P.C., Demozay, D., Hall, B.A., Li, M., Husain, A., Wright, A.J., Atkinson, M.A., and Rhodes, C.J. (2012). Formation of a human beta-cell population within pancreatic islets is set early in life. J Clin Endocrinol Metab 97, 3197-3206.

Gromada, J., Brock, B., Schmitz, O., and Rorsman, P. (2004). Glucagon-Like Peptide-1: Regulation of Insulin Secretion and Therapeutic Potential. Basic and Clinical Pharmacology and Toxicology 95, 252-262.

Gromada, J., Chabosseau, P., and Rutter, G.A. (2018). The alpha-cell in diabetes mellitus. Nat Rev Endocrinol.

Gromada, J., Franklin, I., and Wollheim, C.B. (2007). Alpha-cells of the endocrine pancreas: 35 years of research but the enigma remains. Endocr Rev 28, 84-116.

Groop, L.C., Bonadonna, R.C., DelPrato, S., Ratheiser, K., Zyck, K., Ferrannini, E., and DeFronzo, R.A. (1989). Glucose and free fatty acid metabolism in non-insulin-dependent diabetes mellitus. Evidence for multiple sites of insulin resistance. J Clin Invest 84, 205-213.

Group, S.S. (2004). SEARCH for Diabetes in Youth: a multicenter study of the prevalence, incidence and classification of diabetes mellitus in youth. Control Clin Trials 25, 458-471.

Gu, G., Dubauskaite, J., and Melton, D.A. (2002). Direct evidence for the pancreatic lineage: NGN3+ cells are islet progenitors and are distinct from duct progenitors. Development 129, 2447-2457.

Guilherme, A., Virbasius, J.V., Puri, V., and Czech, M.P. (2008). Adipocyte dysfunctions linking obesity to insulin resistance and type 2 diabetes. Nat Rev Mol Cell Biol 9, 367-377.

Gutierrez, G.D., Bender, A.S., Cirulli, V., Mastracci, T.L., Kelly, S.M., Tsirigos, A., Kaestner, K.H., and Sussel, L. (2017). Pancreatic beta cell identity requires continual repression of non-beta cell programs. J Clin Invest 127, 244-259.

Guttman, M., Amit, I., Garber, M., French, C., Lin, M.F., Feldser, D., Huarte, M., Zuk, O., Carey, B.W., Cassady, J.P., et al. (2009). Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature 458, 223-227.

Guttman, M., Donaghey, J., Carey, B.W., Garber, M., Grenier, J.K., Munson, G., Young, G., Lucas, A.B., Ach, R., Bruhn, L., et al. (2011). lincRNAs act in the circuitry controlling pluripotency and differentiation. Nature 477, 295-300.

Guttman, M., and Rinn, J.L. (2012). Modular regulatory principles of large non-coding RNAs. Nature 482, 339-346.

Guz, Y., Montminy, M.R., Stein, R., Leonard, J., Gamer, L.W., Wright, C.V.E., and Teitelman, G. (1995). Expression of murine STF-1, a putative insulin gene transcription factor, in β cells of pancreas, duodenal epithelium and pancreatic exocrine and endocrine progenitors during ontogeny. Development 121, 11-18.

Hahn, M.W. and Wray, G.A. (2002). The g-value paradox. Evol Dev *4*, 73-75.

Hancock, A.S., Du, A., Liu, J., Miller, M., and May, C.L. (2010). Glucagon deficiency reduces hepatic glucose production and improves glucose tolerance in adult mice. Mol Endocrinol 24, 1605-1614.

Hang, Y., and Stein, R. (2011). MafA and MafB activity in pancreatic beta cells. Trends Endocrinol Metab 22, 364-373.

Hanson, R.L., Craig, D.W., Millis, M.P., Yeatts, K.A., Kobes, S., Pearson, J.V., Lee, A.M., Knowler, W.C., Nelson, R.G., and Wolford, J.K. (2007). Identification of PVT1 as a candidate gene for end-stage renal disease in type 2 diabetes using a pooling-based genome-wide single nucleotide polymorphism association study. Diabetes 56, 975-983.

Hart, A.W., Mella, S., Mendrychowski, J., van Heyningen, V., and Kleinjan, D.A. (2013). The developmental regulator Pax6 is essential for maintenance of islet cell function in the adult mouse pancreas. PLoS One 8, e54173.

Hauge-Evans, A.C., King, A.J., Carmignac, D., Richardson, C.C., Robinson, I.C., Low, M.J., Christie, M.R., Persaud, S.J., and Jones, P.M. (2009). Somatostatin secreted by islet delta-cells fulfills multiple roles as a paracrine regulator of islet function. Diabetes 58, 403-411.

Hayashi, Y., Yamamoto, M., Mizoguchi, H., Watanabe, C., Ito, R., Yamamoto, S., Sun, X.Y., and Murata, Y. (2009). Mice deficient for glucagon gene-derived peptides display normoglycemia and hyperplasia of islet {alpha}-cells but not of intestinal L-cells. Mol Endocrinol 23, 1990-1999.

Hebrok, M. (2012). Generating beta cells from stem cells-the story so far. Cold Spring Harb Perspect Med 2, a007674.

Hebrok, M., Kim, S.K., St-Jacques, B., McMahon, A.P., and Melton, D.A. (2000). Regulation of pancreas development by hedgehog signaling. Development 127, 4905-4913.

Heimberg, H., De Vos, A., Pipeleers, D., Thorens, B., and Schuit, F. (1995). Differences in Glucose Transporter Gene Expression between Rat Pancreatic alpha and beta-cells are

Correlated to Differences in Glucose Transport but Not in Glucose Utilization. The Journal of Biological Chemistry 270, 8971-8975.

Heller, R.S., Stoffers, D.A., Liu, A., Schedl, A., Crenshaw, E.B., 3rd, Madsen, O.D., and Serup, P. (2004). The role of Brn4/Pou3f4 and Pax6 in forming the pancreatic glucagon cell identity. Dev Biol 268, 123-134.

Henseleit, K.D., Nelson, S.B., Kuhlbrodt, K., Hennings, J.C., Ericson, J., and Sander, M. (2005). NKX6 transcription factor activity is required for alpha- and beta-cell development in the pancreas. Development 132, 3139-3149.

Holland, A.M., Hale, M.A., Kagami, H., Hammer, R.E., and MacDonald, R.J. (2002). Experimental control of pancreatic development and maintenance. Proc Natl Acad Sci U S A 99, 12236-12241.

Holman, R.R., Sourij, H., and Califf, R.M. (2014). Cardiovascular outcome trials of glucose-lowering drugs or strategies in type 2 diabetes. The Lancet 383, 2008-2017.

Holst, J.J., Wewer Albrechtsen, N.J., Pedersen, J., and Knop, F.K. (2017). Glucagon and Amino Acids Are Linked in a Mutual Feedback Cycle: The Liver-alpha-Cell Axis. Diabetes 66, 235-240.

Hou, J., Li, Z., Zhong, W., Hao, Q., Lei, L., Wang, L., Zhao, D., Xu, P., Zhou, Y., Wang, Y., et al. (2017). Temporal Transcriptomic and Proteomic Landscapes of Deteriorating Pancreatic Islets in Type 2 Diabetic Rats. Diabetes 66, 2188-2200.

Huang, H.P., Liu, M., El-Hodiri, H.M., Chu, K., Jamrich, M., and Tsai, M.J. (2000). Regulation of the Pancreatic Islet-Specific Gene BETA2 (neuroD) by Neurogenin 3. Molecular and Cellular Biology 20, 3292-3307.

International Diabetes Federation. (2017). IDF Diabetes Atlas, 8th edition. Brussels, Belgium. http://www.diabetesatlas.org.

Inzucchi, S.E., Umpierrez, G., DiGenio, A., Zhou, R., and Kovatchev, B. (2015). How well do glucose variability measures predict patient glycaemic outcomes during treatment intensification in type 2 diabetes? Diabetes Research and Clinical Practice 110, 234-240.

Ishihara, H., Maechler, P., Gjinovci, A., Herrera, P.L., and Wollheim, C.B. (2003). Islet beta-cell secretion determines glucagon release from neighbouring alpha-cells. Nat Cell Biol 5, 330-335.

Jacquemin, P., Lemaigre, F.P., and Rousseau, G.G. (2003). The Onecut transcription factor HNF-6 (OC-1) is required for timely specification of the pancreas and acts upstream of Pdx-1 in the specification cascade. Developmental Biology 258, 105-116.

Jensen, J., Heller, R.S., Funder-Nielsen, T., Pedersen, E.E., Lindsell, C., Weinmaster, G., Madsen, O.D., and Serup, P. (2000). Independent development of pancreatic and β-cells from Neurogenin3- expressing precursors: A role for the notch pathway in repression of premature differentiation. Diabetes 49, 163-176.

Jeon, C.Y., Lokken, R.P., Hu, F.B., and van Dam, R.M. (2007). Physical activity of moderate intensity and risk of type 2 diabetes: a systematic review. Diabetes Care 30, 744-752.

Jiang, W., Liu, Y., Liu, R., Zhang, K., and Zhang, Y. (2015). The lncRNA DEANR1 facilitates human endoderm differentiation by activating FOXA2 expression. Cell Rep 11, 137-148.

Jin, T. (2008). Mechanisms underlying proglucagon gene expression. J Endocrinol 198, 17-28.

Joglekar, M.V., Patil, D., Joglekar, V.M., Rao, G.V., Reddy, D.N., Mitnala, S., Shouche, Y., and Hardikar, A.A. (2009). The miR-30 family microRNAs confer epithelial phenotype to human pancreatic cells. Islets 1, 137-147.

Jonkers, F.C., and Henquin, J.C. (2001). Measurements of Cytoplasmic Ca2. Diabetes 50, 540-550.

Jorgensen, M.C., Ahnfelt-Ronne, J., Hald, J., Madsen, O.D., Serup, P., and Hecksher-Sorensen, J. (2007). An illustrated review of early pancreas development in the mouse. Endocr Rev 28, 685-705.

Juan-Mateu, J., Villate, O., Eizirik, D.L. (2016). Alternative splicing: the new frontier in diabetes research. European Journal of Endocrinology *174*, R225–R238.

Kalis, M., Bolmeson, C., Esguerra, J.L., Gupta, S., Edlund, A., Tormo-Badia, N., Speidel, D., Holmberg, D., Mayans, S., Khoo, N.K., et al. (2011). Beta-cell specific deletion of Dicer1 leads to defective insulin secretion and diabetes mellitus. PLoS One 6, e29166.

Kallman, F. (1964). Fine Structure of Differentiating Mouse Pancreatic Exocrine Cells in Transfilter Culture. The Journal of Cell Biology 20, 399-413.

Kanda, T., Sullivan, K.F., and Wahl, G.M. (1998). Histone–GFP fusion protein enables sensitive analysis of chromosome dynamics in living mammalian cells. Current Biology 8, 377-385.

Kang, H.S., Takeda, Y., Jeon, K., and Jetten, A.M. (2016). The Spatiotemporal Pattern of Glis3 Expression Indicates a Regulatory Function in Bipotent and Endocrine Progenitors during Early Pancreatic Development and in Beta, PP and Ductal Cells. PLoS One 11, e0157138.

Kanji, M.S., Martin, M.G., and Bhushan, A. (2013). Dicer1 Is Required to Repress Neuronal Fate During Endocrine Cell Maturation. Diabetes 62, 1602-1611.

Karro, J.E., Yan, Y., Zheng, D., Zhang, Z., Carriero, N., Cayting, P., Harrrison, P., and Gerstein, M. (2007). Pseudogene.org: a comprehensive database and comparison platform for pseudogene annotation. Nucleic Acids Res 35, D55-60.

Kassem, S.A., Ariel, I., Thornton, P.S., Scheimberg, I., and Glaser, B. (2000). Beta. Diabetes 49, 1325-1333.

Katsarou, A., Gudbjornsdottir, S., Rawshani, A., Dabelea, D., Bonifacio, E., Anderson, B.J., Jacobsen, L.M., Schatz, D.A., and Lernmark, A. (2017). Type 1 diabetes mellitus. Nat Rev Dis Primers 3, 17016.

Katsuura, G., Asakawa, A., and Inui, A. (2002). Roles of pancreatic polypeptide in regulation of food intake. Peptides 23, 323-329.

Kawaguchi, Y., Cooper, B., Gannon, M., Ray, M., MacDonald, R.J., and Wright, C.V. (2002). The role of the transcriptional regulator Ptf1a in converting intestinal to pancreatic progenitors. Nat Genet 32, 128-134.

Keenan, H.A., Sun, J.K., Levine, J., Doria, A., Aiello, L.P., Eisenbarth, G., Bonner-Weir, S., and King, G.L. (2010). Residual insulin production and pancreatic ss-cell turnover after 50 years of diabetes: Joslin Medalist Study. Diabetes 59, 2846-2853.

Keller, A., Peltzer, J., Carpentier, G., Horvath, I., Olah, J., Duchesnay, A., Orosz, F., and Ovadi, J. (2007). Interactions of enolase isoforms with tubulin and microtubules during myogenesis. Biochim Biophys Acta 1770, 919-926.

Kennedy, H.J., Pouli, A.E., Ainscow, E.K., Jouaville, L.S., Rizzuto, R., and Rutter, G.A. (1999). Glucose Generates Sub-plasma Membrane ATP Microdomains in Single Islet The Journal of Biological Chemistry 274, 13281-13291.

Khan, A.H., and Pessin, J.E. (2002). Insulin regulation of glucose uptake: a complex interplay of intracellular signalling pathways. Diabetologia 45, 1475-1483.

Kieffer, T.J. (2016). Closing in on Mass Production of Mature Human Beta Cells. Cell Stem Cell 18, 699-702.

Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low memory requirements. Nat Methods 12, 357-360.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol 14, 1-13.

Kim, E., Magen, A., and Ast, G. (2007). Different levels of alternative splicing among eukaryotes. Nucleic Acids Research *35*, 125–131.

Kim, M.S., Pinto, S.M., Getnet, D., Nirujogi, R.S., Manda, S.S., Chaerkady, R., Madugundu, A.K., Kelkar, D.S., Isserlin, R., Jain, S., et al. (2014). A draft map of the human proteome. Nature *509*, 575–581.

Kim, S., Gupta, N., and Pevzner, P.A. (2008). Spectral Probabilities and Generating Functions of Tandem Mass Spectra: A Strike against Decoy Databases. Journal of Proteome Research 7, 3354-3363.

Kim, S., and Pevzner, P.A. (2014). MS-GF+ makes progress towards a universal database search tool for proteomics. Nat Commun 5, 5277.

Kimball, C.P., and Murlin, J.R. (1923). Aqueous Extracts of Pancreas. American Journal of Physiology 58, 337-346.

Kino, T., Hurt, D.E., Ichijo, T., Nader, N., and Chrousos, G.P. (2010). Noncoding RNA Gas5 Is a Growth Arrest– and Starvation-Associated Repressor of the Glucocorticoid Receptor. Molecular Biology 3, 1-16.

Kiselev, Y., Eriksen, T.E., Forsdahl, S., Nguyen, L.H., and Mikkola, I. (2012). 3T3 cell lines stably expressing Pax6 or Pax6(5a)--a new tool used for identification of common and isoform specific target genes. PLoS One 7, e31915.

Kleiner, I. (1919). The action of intravenous injections of pancreas emulsions in experimental diabetes. J. Biol. Chem. 40, 153–170.

Kleinridders, A., Ferris, H.A., Cai, W., and Kahn, C.R. (2014). Insulin action in brain regulates systemic metabolism and brain function. Diabetes 63, 2232-2243.

Klinck, R., Fuchtbauer, E.M., Ahnfelt-Ronne, J., Serup, P., Jensen, J.N., and Jorgensen, M.C. (2011). A BAC transgenic Hes1-EGFP reporter reveals novel expression domains in mouse embryos. Gene Expr Patterns 11, 415-426.

Kohn, A.M., Summers, S.A., Birnbaum, M.J., and Roth, R.A. (1996). Expression of a Constitutively Active Akt Ser/Thr Kinase in 3T3-L1 Adipocytes Stimulates Glucose Uptake and Glucose Transporter 4 Translocation. The Journal of Biological Chemistry 271, 31372-31378.

Komatsu, M., Takei, M., Ishii, H., and Sato, Y. (2013). Glucose-stimulated insulin secretion: A newer perspective. J Diabetes Investig 4, 511-516.

Kong, L., Zhang, Y., Ye, Z.Q., Liu, X.Q., Zhao, S.Q., Wei, L., and Gao, G. (2007). CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. Nucleic Acids Res 35, W345-349.

Kopinke, D., Brailsford, M., Pan, F.C., Magnuson, M.A., Wright, C.V., and Murtaugh, L.C. (2012). Ongoing Notch signaling maintains phenotypic fidelity in the adult exocrine pancreas. Dev Biol 362, 57-64.

Kopp, J.L., Dubois, C.L., Schaffer, A.E., Hao, E., Shih, H.P., Seymour, P.A., Ma, J., and Sander, M. (2011). Sox9+ ductal cells are multipotent progenitors throughout development but do not produce new endocrine cells in the normal or injured adult pancreas. Development 138, 653-665.

Kovarovic, K., and Andrews, P. (2007). Bovid postcranial ecomorphological survey of the Laetoli paleoenvironment. J Hum Evol 52, 663-680.

Krapp, A., Knofler, M., Ledermann, B., Burki, K., Berney, C., Zoerkler, N., Hagenbuchle, O., and Wellauer, P.K. (1998). The bHLH protein PTF1-p48 is essential for the formation of the

exocrine and the correct spatial organization of the endocrine pancreas. Genes and Development 12, 3752-3763.

Kredo-Russo, S., Mandelbaum, A.D., Ness, A., Alon, I., Lennox, K.A., Behlke, M.A., and Hornstein, E. (2012). Pancreas-enriched miRNA refines endocrine cell differentiation. Development 139, 3021-3031.

Kroon, E., Martinson, L.A., Kadoya, K., Bang, A.G., Kelly, O.G., Eliazer, S., Young, H., Richardson, M., Smart, N.G., Cunningham, J., et al. (2008). Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells in vivo. Nat Biotechnol 26, 443-452.

Ku, G.M., Kim, H., Vaughn, I.W., Hangauer, M.J., Myung Oh, C., German, M.S., and McManus, M.T. (2012). Research resource: RNA-Seq reveals unique features of the pancreatic beta-cell transcriptome. Mol Endocrinol 26, 1783-1792.

Kuhn, R.M., Haussler, D., and Kent, W.J. (2013). The UCSC genome browser and associated tools. Brief Bioinform 14, 144-161.

Kulkarni, R.N. (2005). New Insights into the Roles of Insulin/IGF-I in the Development and Maintenance of β-Cell Mass. Reviews in Endocrine & Metabolic Disorders 6, 199-210.

Kushner, Jake A., MacDonald, Patrick E., and Atkinson, Mark A. (2014). Stem Cells to Insulin Secreting Cells: Two Steps Forward and Now a Time to Pause? Cell Stem Cell 15, 535-536.

Lacy, P.E., and Scharp, D.W. (1986). Islet Transplantation In Treating Diabetes. Annual Review Medicine 37, 33-40.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al.; International Human Genome Sequencing Consortium (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860–921.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods 9, 357-359.

LaPierre, M.P., and Stoffel, M. (2017). MicroRNAs as stress regulators in pancreatic beta cells and diabetes. Mol Metab 6, 1010-1023.

Larsen, H.L., and Grapin-Botton, A. (2017). The molecular and morphogenetic basis of pancreas organogenesis. Semin Cell Dev Biol 66, 51-68.

Latreille, M., Hausser, J., Stutzer, I., Zhang, Q., Hastoy, B., Gargani, S., Kerr-Conte, J., Pattou, F., Zavolan, M., Esguerra, J.L., et al. (2014). MicroRNA-7a regulates pancreatic beta cell function. J Clin Invest 124, 2722-2735.

Lee, J.C., Smith, S.B., Watada, H., Lin, J., Scheel, D., Wang, J., Mirmira, R.G., and German, M.S. (2001). Regulation of the Pancreatic Pro-Endocrine Gene Neurogenin3. Diabetes 50, 928-936.

Li, B., Bi, C.L., Lang, N., Li, Y.Z., Xu, C., Zhang, Y.Q., Zhai, A.X., and Cheng, Z.F. (2014). RNA-seq methods for identifying differentially expressed gene in human pancreatic islet cells treated with pro-inflammatory cytokines. Mol Biol Rep 41, 1917-1925.

Li, B., Qing, T., Zhu, J., Wen, Z., Yu, Y., Fukumura, R., Zheng, Y., Gondo, Y., and Shi, L. (2017). A Comprehensive Mouse Transcriptomic BodyMap across 17 Tissues by RNA-seq. Sci Rep 7, 4200.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078-2079.

Li, W., Cavelti-Weder, C., Zhang, Y., Clement, K., Donovan, S., Gonzalez, G., Zhu, J., Stemann, M., Xu, K., Hashimoto, T., et al. (2014). Long-term persistence and development of induced pancreatic beta cells generated by lineage conversion of acinar cells. Nat Biotechnol 32, 1223-1230.

Like, A.A. (1967). The ultrastructure of the secretory cells of the islets of Langerhans in man. Lab Invest. *6*, 937-951.

Lin, M.F., Jungreis, I., and Kellis, M. (2011). PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions. Bioinformatics 27, i275-282.

Lin, S., Lin, Y., Nery, J.R., Urich, M.A., Breschi, A., Davis, C.A., Dobin, A., Zaleski, C., Beer, M.A., Chapman, W.C., et al. (2014). Comparison of the transcriptional landscapes between human and mouse tissues. Proc Natl Acad Sci U S A 111, 17224-17229.

Longuet, C., Robledo, A.M., Dean, E.D., Dai, C., Ali, S., McGuinness, I., de Chavez, V., Buguin, P.M., Charron, M.J., Powers, A.C., et al. (2013). Liver-Specific Disruption of the Murine Glucagon Receptor Produces a-Cell Hyperplasia, Evidence for a Circulating Alpha Cell Growth Factor. Diabetes 62, 1196-1205.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15, 550.

Lowry, W.E., and Plath, K. (2008). The many ways to make an iPS cell. Nature Biotechnology 26, 1246-1248.

Lu, T.T., Heyne, S., Dror, E., Casas, E., Leonhardt, L., Boenke, T., Yang, C.H., Sagar, Arrigoni, L., Dalgaard, K., et al. (2018). The Polycomb-Dependent Epigenome Controls beta Cell Dysfunction, Dedifferentiation, and Diabetes. Cell Metab 27, 1294-1308 e1297.

Lubeck, E., and Cai, L. (2012). Single-cell systems biology by super-resolution imaging and combinatorial labeling. Nat Methods 9, 743-748.

Luo, H., Bu, D., Sun, L., Fang, S., Liu, Z., and Zhao, Y. (2017). Identification and function annotation of long intervening noncoding RNAs. Brief Bioinform 18, 789-797.

Lynn, F.C., Skewes-Cox, P., Kosaka, Y., McManus, M.T., Harfe, B.D., and German, M.S. (2007). MicroRNA expression is required for pancreatic islet cell genesis in the mouse. Diabetes 56, 2938-2945.

MacCuish AC, Munro JF, Duncan LJ. Treatment of hypoglycaemic coma with gluca- gon, intravenous dextrose, and mannitol infusion in a hundred diabetics. Lancet 2: 946 –949, 1970.

MacDonald, P.E., El-kholy, W., Riedel, M.J., Salapatek, A.N.F., Light, P.E., and Wheeler, M.B. (2002). The Multiple Actions of GLP-1 on the Process of Glucose-Stimulated Insulin Secretion. Diabetes 51, S434-S442.

Magnusson, I., Rothman, D.L., Katz, L.D., Shulman, R.G., and Shulman, G.I. (1992). Increased rate of gluconeogenesis in type II diabetes mellitus. A 13C nuclear magnetic resonance study. J Clin Invest 90, 1323-1327.

Mandelbaum, A.D., Melkman-Zehavi, T., Oren, R., Kredo-Russo, S., Nir, T., Dor, Y., and Hornstein, E. (2012). Dysregulation of Dicer1 in beta cells impairs islet architecture and glucose metabolism. Exp Diabetes Res 2012, 470302.

Manuel, M.N., Mi, D., Mason, J.O., and Price, D.J. (2015). Regulation of cerebral cortical neurogenesis by the Pax6 transcription factor. Front Cell Neurosci 9, 70.

Marty-Santos, L., and Cleaver, O. (2016). Pdx1 regulates pancreas tubulogenesis and E-cadherin expression. Development 143, 1056.

Maruyama, H., Tominaga, M., Bolli, G., Orci, L., and Unger, R.H. (1985). The alpha cell response to glucose change during perfusion of anti-insulin serum in pancreas isolated from normal rats. Diabetologia 28, 836-840.

Matsuoka, T.A. et al. (2004) The MafA transcription factor appears to be responsible for tissue-specific expression of insulin. Proc. Natl. Acad. Sci. U.S.A. 101, 2930–2933.

Matsuoka, T., Kawashima, S., Miyatsuka, T., Sasaki, S., Shimo, N., Katakami, N., Kawamori, D., Takebe, S., Herrera, P.L., Kaneto, H., et al. (2017). Mafa Enables Pdx1 to Effectively Convert Pancreatic Islet Progenitors and Committed Islet a-Cells Into b-Cells In Vivo. diabetes 66, 1293-1300.

Maurano, M.T., Humbert, R., Rynes, E., Thurman, R.E., Haugen, E., Wang, H., Reynolds, A.P., Sandstrom, R., Qu, H., Brody, J., et al. (2012). Systematic Localization of Common Disease-Associated Variation in Regulatory DNA. Science 337, 1190-1195.

Mayampurath, A.M., Jaitly, N., Purvine, S.O., Monroe, M.E., Auberry, K.J., Adkins, J.N., and Smith, R.D. (2008). DeconMSn: a software tool for accurate parent ion monoisotopic mass determination for tandem mass spectra. Bioinformatics 24, 1021-1023.

McCarthy, M.I. (2010). Genomics, Type 2 Diabetes, and Obesity. Genomic Medicine 363, 2339-2350.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol 28, 495-501.

Meier, J.J., Butler, A.E., Saisho, Y., Monchamp, T., Galasso, R., Bhushan, A., Rizza, R.A., and Butler, P.C. (2008). Beta-cell replication is the primary mechanism subserving the postnatal expansion of beta-cell mass in humans. Diabetes 57, 1584-1594.

Melkman-Zehavi, T., Oren, R., Kredo-Russo, S., Shapira, T., Mandelbaum, A.D., Rivkin, N., Nir, T., Lennox, K.A., Behlke, M.A., Dor, Y., et al. (2011). miRNAs control insulin content in pancreatic beta-cells via downregulation of transcriptional repressors. EMBO J 30, 835-845.

Mellacheruvu, D., Wright, Z., Couzens, A.L., Lambert, J.P., St-Denis, N.A., Li, T., Miteva, Y.V., Hauri, S., Sardiu, M.E., Low, T.Y., et al. (2013). The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. Nat Methods 10, 730-736.

Mellitzer, G., Bonne, S., Luco, R.F., Van de Casteele, M., Lenne-Samuel, N., Collombat, P., Mansouri, A., Lee, J., Lan, M., Pipeleers, D., et al. (2005). IA1 is NGN3-dependent and essential for differentiation of the endocrine pancreas. The EMBO Journal 25, 1344-1352.

Mendell, J.T., and Olson, E.N. (2012). MicroRNAs in stress signaling and human disease. Cell 148, 1172-1187.

Menendez-Gutierrez, M.P., Roszer, T., Fuentes, L., Nunez, V., Escolano, A., Redondo, J.M., De Clerck, N., Metzger, D., Valledor, A.F., and Ricote, M. (2015). Retinoid X receptors orchestrate osteoclast differentiation and postnatal bone remodeling. J Clin Invest 125, 809-823.

Mercer, T.R., Dinger, M.E., and Mattick, J.S. (2009). Long non-coding RNAs: insights into functions. Nature Reviews Genetics 10, 155-159.

Mezza, T., and Kulkarni, R.N. (2014). The regulation of pre- and post-maturational plasticity of mammalian islet cell mass. Diabetologia 57, 1291-1303.

Mikami, S.I., Ono, K. (1962). Glucagon deficiency induced by extirpation of alpha islets of the fowl pancreas. Endocrinology. *71*, 464-473

Millman, J.R., and Pagliuca, F.W. (2017). Autologous Pluripotent Stem Cell-Derived beta-Like Cells for Diabetes Cellular Therapy. Diabetes 66, 1111-1120.

Mitchell, R.K., Nguyen-Tu, M.S., Chabosseau, P., Callingham, R.M., Pullen, T.J., Cheung, R., Leclerc, I., Hodson, D.J., and Rutter, G.A. (2017). The transcription factor Pax6 is required for pancreatic beta cell identity, glucose-regulated ATP synthesis, and Ca(2+) dynamics in adult mice. J Biol Chem 292, 8892-8906.

Miyatsuka, T., Kosaka, Y., Kim, H., and German, M.S. (2011). Neurogenin3 inhibits proliferation in endocrine progenitors by inducing Cdkn1a. Proc Natl Acad Sci U S A 108, 185-190.

Miyazaki, J.I., Araki, K., Yamato, E., Ikegami, H., Asano, T., Shibasaki, Y., Oka, Y., and Yamamura, K.I. (1990). Establishment of a pancreatic beta cell line that retains glucose-inducible insulin secretion: special reference to expression of glucose transporter isoforms. Endocrinology 127, 126-132.

Molven, A., Hollister-Lock, J., Hu, J., Martinez, R., Njolstad, P.R., Liew, C.W., Weir, G., and Kulkarni, R.N. (2016). The Hypoglycemic Phenotype Is Islet Cell-Autonomous in Short-Chain Hydroxyacyl-CoA Dehydrogenase-Deficient Mice. Diabetes 65, 1672-1678.

Moran, I., Akerman, I., van de Bunt, M., Xie, R., Benazra, M., Nammo, T., Arnes, L., Nakic, N., Garcia-Hurtado, J., Rodriguez-Segui, S., et al. (2012). Human beta cell transcriptome analysis uncovers lncRNAs that are tissue-specific, dynamically regulated, and abnormally expressed in type 2 diabetes. Cell Metab 16, 435-448.

Motterle, A., Gattesco, S., Caille, D., Meda, P., and Regazzi, R. (2015). Involvement of long non-coding RNAs in beta cell failure at the onset of type 1 diabetes in NOD mice. Diabetologia 58, 1827-1835.

Motterle, A., Gattesco, S., Peyot, M.L., Esguerra, J.L.S., Gomez-Ruiz, A., Laybutt, D.R., Gilon, P., Burdet, F., Ibberson, M., Eliasson, L., et al. (2017). Identification of islet-enriched long non-coding RNAs contributing to beta-cell failure in type 2 diabetes. Mol Metab 6, 1407-1418.

Motterle, A., Sanchez-Parra, C., and Regazzi, R. (2016). Role of long non-coding RNAs in the determination of β-cell identity. Diabetes Obes Metab 18, 41-50.

Mullen, Y.S., Clark, W.R., Molnar, I.G., and Brown, J. (1977). Complete Reversal of Experimental Diabetes Mellitus in Rats by a Single Fetal Pancreas. Science 195, 68-70.

Muller, T.D., Finan, B., Clemmensen, C., DiMarchi, R.D., and Tschop, M.H. (2017). The New Biology and Pharmacology of Glucagon. Physiol Rev 97, 721-766.

Munger, B.L. (1958). A light and electron microscopic study of cellular differentiation in the pancreatic islets of the mouse. American Journal of Anatomy 103, 275-311.

Murtaugh, L.C. (2007). Pancreas and beta-cell development: from the actual to the possible. Development 134, 427-438.

Nagasaki, H., Arita, M., Nishizawa, T., Suwa, M., Gotoh, O. (2005). Species- specific variation of alternative splicing and transcriptional initiation in six eukaryotes. Gene *364, 53-62.*

Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., Okita, K., Mochiduki, Y., Takizawa, N., and Yamanaka, S. (2008). Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. Nat Biotechnol 26, 101-106.

Nam, H.S., and Benezra, R. (2009). High levels of Id1 expression define B1 type adult neural stem cells. Cell Stem Cell 5, 515-526.

Naya, F.J., Huang, H.P., Qiu, Y., Mutoh, H., DeMayo, F.J., Leiter, A.B., and Tsai, M.J. (1997). Diabetes, defective pancreatic morphogenesis, and abnormal enteroendocrine differentiation in BETA2/NeuroD-deficient mice. Genes and Development 11, 2323-2334.

Nica, A.C., Ongen, H., Irminger, J.C., Bosco, D., Berney, T., Antonarakis, S.E., Halban, P.A., and Dermitzakis, E.T. (2013). Cell-type, allelic, and genetic signatures in the human pancreatic beta cell transcriptome. Genome Res 23, 1554-1562.

Nir, T., Melton, D.A., and Dor, Y. (2007). Recovery from diabetes in mice by beta cell regeneration. J Clin Invest 117, 2553-2561.

Nishimura, W., Kondo, T., Salameh, T., El Khattabi, I., Dodge, R., Bonner-Weir, S., and Sharma, A. (2006). A switch from MafB to MafA expression accompanies differentiation to pancreatic beta-cells. Dev Biol 293, 526-539.

Nostro, M.C., and Keller, G. (2012). Generation of beta cells from human pluripotent stem cells: Potential for regenerative medicine. Semin Cell Dev Biol 23, 701-710.

Obata, J., Yano, M., Mimura, H., Goto, T., Nakayama, R., Mibu, Y., Oka, C., and Kawaichi, M. (2001). p48 subunit of mouse PTF1 binds to RBP-Jk/CBF-1, the intracellular mediator of Notch signalling, and is expressed in the neural tube of early stage embryos. Genes to Cells 2001, 345-360.

Offield, M.F., Jetton, T.L., Labosky, P.A., Ray, M., Stein, R.W., Magnuson, M.A., Hogan, B.L.M., and Wright, C.V.E. (1996). PDX-1 is required for pancreatic outgrowth and differentiation of the rostral duodenum. Development 122, 983-995.

Oliver-Krasinski, J.M., and Stoffers, D.A. (2008). On the origin of the beta cell. Genes Dev 22, 1998-2021.

Opie, E. (1900). On the relation of chronic interstitial pancreatitis to the islands of Langerhans and to diabetes mellitus. J. Exp. Med. 5, 419–428.

Oram, R.A., McDonald, T.J., Shields, B.M., Hudson, M.M., Shepherd, M.H., Hammersley, S., Pearson, E.R., and Hattersley, A.T. (2015). Most People With Long-Duration Type 1 Diabetes in a Large Population-Based Study Are Insulin Microsecretors. Diabetes Care 38, 323-328.

Orci, L., and Unger, R.H. (1975). Functional Subdivision of Islets of Langerhans and Possible Role of D Cells. The Lancet 1, 1243-1244.

Pagliuca, F.W., and Melton, D.A. (2013). How to make a functional beta-cell. Development 140, 2472-2483.

Pagliuca, F.W., Millman, J.R., Gurtler, M., Segel, M., Van Dervort, A., Ryu, J.H., Peterson, Q.P., Greiner, D., and Melton, D.A. (2014). Generation of functional human pancreatic beta cells in vitro. Cell 159, 428-439.

Pan, F.C., and Wright, C. (2011). Pancreas organogenesis: from bud to plexus to gland. Dev Dyn 240, 530-565.

Papizan, J.B., Singer, R.A., Tschen, S.I., Dhawan, S., Friel, J.M., Hipkens, S.B., Magnuson, M.A., Bhushan, A., and Sussel, L. (2011). Nkx2.2 repressor complex regulates islet beta-cell specification and prevents beta-to-alpha-cell reprogramming. Genes Dev 25, 2291-2305.

Paulescu N.C. (1920). Syndromes Pancréatiques: Diabetes, Physiologie Médicale, Bucarest, 295–327.

Pavlaki, I., Alammari, F., Sun, B., Clark, N., Sirey, T., Lee, S., Woodcock, D.J., Ponting, C.P., Szele, F.G., and Vance, K.W. (2018). The long non-coding RNA Paupar promotes KAP1-dependent chromatin changes and regulates olfactory bulb neurogenesis. EMBO J 37.

Paz, I., Kosti, I., Ares, M., Jr., Cline, M., and Mandel-Gutfreund, Y. (2014). RBPmap: a web server for mapping binding sites of RNA-binding proteins. Nucleic Acids Res 42, W361-367.

Pedersen, H.K., Gudmundsdottir, V., and Brunak, S. (2017). Pancreatic Islet Protein Complexes and Their Dysregulation in Type 2 Diabetes. Front Genet 8, 43.

Pefanis, E., Wang, J., Rothschild, G., Lim, J., Kazadi, D., Sun, J., Federation, A., Chao, J., Elliott, O., Liu, Z.P., et al. (2015). RNA exosome-regulated long non-coding RNA transcription controls super-enhancer activity. Cell 161, 774-789.

Ponting, C.P., Oliver, P.L., and Reik, W. (2009). Evolution and functions of long noncoding RNAs. Cell 136, 629-641.

Porte, D., Jr., and Pupo, A.A. (1969). Insulin responses to glucose: evidence for a two pool system in man. J Clin Invest 48, 2309-2319.

Posselt, A.M., Bellin, M.D., Tavakol, M., Szot, G.L., Frassetto, L.A., Masharani, U., Kerlan, R.K., Fong, L., Vincenti, F.G., Hering, B.J., et al. (2010). Islet transplantation in type 1 diabetics using an immunosuppressive protocol based on the anti-LFA-1 antibody efalizumab. Am J Transplant 10, 1870-1880.

Poy, M.N., Eliasson, L., Krutzfeldt, J., Kuwajima, S., Ma, X., MacDonald, P.E., Pfeffer, S., Tuschl, T., Rajewsky, N., Rorsman, P., et al. (2004). A pancreatic islet-specific microRNA regulates insulin secretion. Nature 432, 226-230.

Poy, M.N., Hausser, J., Trajkovski, M., Braun, M., Collins, S., Rorsman, P., Zavolan, M., and Stoffel, M. (2009). miR-375 maintains normal pancreatic alpha- and beta-cell mass. Proc Natl Acad Sci U S A 106, 5813-5818.

Pozzilli, P., Leslie, R.D., Chan, J., De Fronzo, R., Monnier, L., Raz, I., and Del Prato, S. (2010). The A1C and ABCD of glycaemia management in type 2 diabetes: a physician's personalized approach. Diabetes Metab Res Rev 26, 239-244.

Prado, C.L., Pugh-Bernard, A.E., Elghazi, L., Sosa-Pineda, B., and Sussel, L. (2004). Ghrelin cells replace insulin-producing beta cells in two mouse models of pancreas development. Proc Natl Acad Sci U S A 101, 2924-2929.

Puri, S., Akiyama, H., and Hebrok, M. (2013). VHL-mediated disruption of Sox9 activity compromises beta-cell identity and results in diabetes mellitus. Genes Dev 27, 2563-2575.

Qu, X., Afelik, S., Jensen, J.N., Bukys, M.A., Kobberup, S., Schmerr, M., Xiao, F., Nyeng, P., Veronica Albertoni, M., Grapin-Botton, A., et al. (2013). Notch-mediated post-translational control of Ngn3 protein stability regulates pancreatic patterning and cell fate commitment. Dev Biol 376, 1-12.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841-842.

Raj, A., van den Bogaard, P., Rifkin, S.A., van Oudenaarden, A., and Tyagi, S. (2008). Imaging individual mRNA molecules using multiple singly labeled probes. Nat Methods 5, 877-879.

Ramalho-Santos, M., Melton, D.A., and McMahon, A.P. (2000). Hedgehog signals regulate multiple aspects of gastrointestinal development. Development 127, 2763-2772.

Ramracheya, R., Ward, C., Shigeto, M., Walker, J.N., Amisten, S., Zhang, Q., Johnson, P.R., Rorsman, P., and Braun, M. (2010). Membrane potential-dependent inactivation of voltage-gated ion channels in alpha-cells inhibits glucagon secretion from human islets. Diabetes 59, 2198-2208.

Raum, J.C., Gerrish, K., Artner, I., Henderson, E., Guo, M., Sussel, L., Schisler, J.C., Newgard, C.B., and Stein, R. (2006). FoxA2, Nkx2.2, and PDX-1 regulate islet beta-cell-specific mafA expression through conserved sequences located between base pairs -8118 and -7750 upstream from the transcription start site. Mol Cell Biol 26, 5735-5743.

Rezania, A., Bruin, J.E., Arora, P., Rubin, A., Batushansky, I., Asadi, A., O'Dwyer, S., Quiskamp, N., Mojibian, M., Albrecht, T., et al. (2014). Reversal of diabetes with insulin-producing cells derived in vitro from human pluripotent stem cells. Nat Biotechnol 32, 1121-1133.

Riedel, M.J., Asadi, A., Wang, R., Ao, Z., Warnock, G.L., and Kieffer, T.J. (2012). Immunohistochemical characterisation of cells co-producing insulin and glucagon in the developing human pancreas. Diabetologia 55, 372-381.

Rinn, J.L., and Chang, H.Y. (2012). Genome regulation by long noncoding RNAs. Annu Rev Biochem 81, 145-166.

Rinn, J.L., Kertesz, M., Wang, J.K., Squazzo, S.L., Xu, X., Brugmann, S.A., Goodnough, L.H., Helms, J.A., Farnham, P.J., Segal, E., et al. (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell 129, 1311-1323.

Ritz-Laser, B., Estreicher, A., Gauthier, B.R., Mamin, A., Edlund, A., and Philippe, J. (2002). The pancreatic beta-cell-specific transcription factor Pax-4 inhibits glucagon gene expression through Pax-6. Diabetologia 45, 97-107.

Roder, P.V., Wu, B., Liu, Y., and Han, W. (2016). Pancreatic regulation of glucose homeostasis. Exp Mol Med 48, e219.

Roman, T.S., Cannon, M.E., Vadlamudi, S., Buchkovich, M.L., Wolford, B.N., Welch, R.P., Morken, M.A., Kwon, G.J., Varshney, A., Kursawe, R., et al. (2017). A Type 2 Diabetes-Associated Functional Regulatory Variant in a Pancreatic Islet Enhancer at the ADCY5 Locus. Diabetes 66, 2521-2530.

Romero-Barrios, N., Legascue, M.F., Benhamed, M., Ariel, F., and Crespi, M. (2018). Splicing regulation by long noncoding RNAs. Nucleic Acids Res 46, 2169-2184.

Rorsman, P., and Braun, M. (2013). Regulation of insulin secretion in human pancreatic islets. Annu Rev Physiol 75, 155-179.

Rosenfeld, L. (2002). Insulin: Discovery and Controversy. Clinical Chemistry 48, 2270-2288.

Rukstalis, J.M., and Habener, J.F. (2007). Snail2, a mediator of epithelial-mesenchymal transitions, expressed in progenitor cells of the developing endocrine pancreas. Gene Expr Patterns 7, 471-479.

Russ, H.A., Parent, A.V., Ringler, J.J., Hennings, T.G., Nair, G.G., Shveygert, M., Guo, T., Puri, S., Haataja, L., Cirulli, V., et al. (2015). Controlled induction of human pancreatic progenitors produces functional beta-like cells in vitro. EMBO J 34, 1759-1772.

Rusu, V., Hoch, E., Mercader, J.M., Tenen, D.E., Gymrek, M., Hartigan, C.R., DeRan, M., von Grotthuss, M., Fontanillas, P., Spooner, A., et al. (2017). Type 2 Diabetes Variants Disrupt Function of SLC16A11 through Two Distinct Mechanisms. Cell 170, 199-212 e120.

Samuel, V.T., and Shulman, G.I. (2012). Mechanisms for insulin resistance: common threads and missing links. Cell 148, 852-871.

Sancho, R., Gruber, R., Gu, G., and Behrens, A. (2014). Loss of Fbw7 reprograms adult pancreatic ductal cells into alpha, delta, and beta cells. Cell Stem Cell 15, 139-153.

Sander, M., Neubuser, A., Kalamaras, J., Ee, H.C., Martin, G.R., and German, M.S. (1997). Genetic analysis reveals that PAX6 is required for normal transcription of pancreatic hormone genes and islet development. Genes and Development 11, 1662-1673.

Sasamoto, Y., Hayashi, R., Park, S.J., Saito-Adachi, M., Suzuki, Y., Kawasaki, S., Quantock, A.J., Nakai, K., Tsujikawa, M., and Nishida, K. (2016). PAX6 Isoforms, along with Reprogramming Factors, Differentially Regulate the Induction of Cornea-specific Genes. Sci Rep 6, 20807.

Saunders, D., and Powers, A.C. (2016). Replicative capacity of beta-cells and type 1 diabetes. J Autoimmun 71, 59-68.

Schaffer, A.E., Freude, K.K., Nelson, S.B., and Sander, M. (2010). Nkx6 transcription factors and Ptf1a function as antagonistic lineage determinants in multipotent pancreatic progenitors. Dev Cell 18, 1022-1029.

Schaffer, A.E., Taylor, B.L., Benthuysen, J.R., Liu, J., Thorel, F., Yuan, W., Jiao, Y., Kaestner, K.H., Herrera, P.L., Magnuson, M.A., et al. (2013). Nkx6.1 controls a gene regulatory network required for establishing and maintaining pancreatic Beta cell identity. PLoS Genet 9, e1003274.

Schulz, N., Liu, K.C., Charbord, J., Mattsson, C.L., Tao, L., Tworus, D., and Andersson, O. (2016). Critical role for adenosine receptor A2a in beta-cell proliferation. Mol Metab 5, 1138-1146.

Schwitzgebel, V.M., Scheel, D.W., Conners, J.R., Kalamaras, J., Lee, J.E., Anderson, D.J., Sussel, L., Johnson, J.D., and German, M.S. (2000). Expression of neurogenin3 reveals an islet cell precursor population in the pancreas. Development 127.

Scott, R.A., Scott, L.J., Magi, R., Marullo, L., Gaulton, K.J., Kaakinen, M., Pervjakova, N., Pers, T.H., Johnson, A.D., Eicher, J.D., et al. (2017). An Expanded Genome-Wide Association Study of Type 2 Diabetes in Europeans. Diabetes 66, 2888-2902.

Scoville, D.W., Cyphert, H.A., Liao, L., Xu, J., Reynolds, A., Guo, S., and Stein, R. (2015). MLL3 and MLL4 Methyltransferases Bind to the MAFA and MAFB Transcription Factors to Regulate Islet b-Cell Function. Diabetes 64, 3772-3783.

Sellick, G.S., Barker, K.T., Stolte-Dijkstra, I., Fleischmann, C., Coleman, R.J., Garrett, C., Gloyn, A.L., Edghill, E.L., Hattersley, A.T., Wellauer, P.K., et al. (2004). Mutations in PTF1A cause pancreatic and cerebellar agenesis. Nat Genet 36, 1301-1305.

Seymour, P.A., Freude, K.K., Tran, M.N., Mayes, E.E., Jensen, J., Kist, R., Scherer, G., and Sander, M. (2007). SOX9 is required for maintenance of the pancreatic progenitor cell pool. Proc Natl Acad Sci U S A 104, 1865-1870.

Seymour, P.A., and Sander, M. (2011). Historical perspective: beginnings of the beta-cell: current perspectives in beta-cell development. Diabetes 60, 364-376.

Shaham, O., Menuchin, Y., Farhy, C., and Ashery-Padan, R. (2012). Pax6: a multi-level regulator of ocular development. Prog Retin Eye Res 31, 351-376.

Shapiro, A.M., Lakey, J.R., Ryan, E.A., Korbutt, G.S., Toth, E., Warnock, G.L., Kneteman, N.M., and Rajotte, R.V. (2000). Islet transplantation in seven patients with type 1 diabetes mellitus using a glucocorticoid-free immunosuppressive regimen. New England Journal of Medicine 343, 230-238.

Shen, S., Park, J.W., Lu, Z.X., Lin, L., Henry, M.D., Wu, Y.N., Zhou, Q., and Xing, Y. (2014). rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. Proc Natl Acad Sci U S A 111, E5593-5601.

Shen, W., Taylor, B., Jin, Q., Nguyen-Tran, V., Meeusen, S., Zhang, Y.Q., Kamireddy, A., Swafford, A., Powers, A.F., Walker, J., et al. (2015). Inhibition of DYRK1A and GSK3B induces human beta-cell proliferation. Nat Commun 6, 8372.

Shih, H.P., Wang, A., and Sander, M. (2013). Pancreas organogenesis: from lineage determination to morphogenesis. Annu Rev Cell Dev Biol 29, 81-105.

Shiota, C., Prasadan, K., Guo, P., El-Gohary, Y., Wiersch, J., Xiao, X., Esni, F., and Gittes, G.K. (2013). alpha-Cells are dispensable in postnatal morphogenesis and maturation of mouse pancreatic islets. Am J Physiol Endocrinol Metab 305, E1030-1040.

Shirakawa, J., Fernandez, M., Takatani, T., El Ouaamari, A., Jungtrakoon, P., Okawa, E.R., Zhang, W., Yi, P., Doria, A., and Kulkarni, R.N. (2017). Insulin Signaling Regulates the FoxM1/PLK1/CENP-A Pathway to Promote Adaptive Pancreatic beta Cell Proliferation. Cell Metab 25, 868-882 e865.

Sigal, R.J., Kenny, G.P., Wasserman, D.H., Castaneda-Sceppa, C., and White, R.D. (2006). Physical activity/exercise and type 2 diabetes: a consensus statement from the American Diabetes Association. Diabetes Care 29, 1433-1438.

Simon, M.D., Wang, C.I., Kharchenko, P.V., West, J.A., Chapman, B.A., Alekseyenko, A.A., Borowsky, M.L., Kuroda, M.I., and Kingston, R.E. (2011). The genomic binding sites of a noncoding RNA. Proc Natl Acad Sci U S A 108, 20497-20502.

Singer, R.A., Arnes, L., and Sussel, L. (2015). Noncoding RNAs in beta cell biology. Curr Opin Endocrinol Diabetes Obes 22, 77-85.

Singer, R.A., and Sussel, L. (2018). Islet Long Noncoding RNAs: A Playbook for Discovery and Characterization. Diabetes 67, 1461-1470.

Sisino, G., Zhou, A.X., Dahr, N., Sabirsh, A., Soundarapandian, M.M., Perera, R., Larsson-Lekholm, E., Magnone, M.C., Althage, M., and Tyrberg, B. (2017). Long noncoding RNAs are dynamically regulated during beta-cell mass expansion in mouse pregnancy and control beta-cell proliferation in vitro. PLoS One 12, e0182371.

Sladek, R., Rocheleau, G., Rung, J., Dina, C., Shen, L., Serre, D., Boutin, P., Vincent, D., Belisle, A., Hadjadj, S., et al. (2007). A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature 445, 881-885.

Smith, M.A., and Mattick, J.S. (2017). Structural and Functional Annotation of Long Noncoding RNAs. Methods Mol Biol 1526, 65-85.

Smith, S.B., Gasa, R., Watada, H., Wang, J., Griffen, S.C., and German, M.S. (2003). Neurogenin3 and Hepatic Nuclear Factor 1 Cooperate in Activating Pancreatic Expression of Pax4. Journal of Biological Chemistry 278, 38254-38259.

Smith, S.B., Watada, H., and German, M.S. (2004). Neurogenin3 activates the islet differentiation program while repressing its own expression. Mol Endocrinol 18, 142-149.

Soccio, R.E., Chen, E.R., Rajapurkar, S.R., Safabakhsh, P., Marinis, J.M., Dispirito, J.R., Emmett, M.J., Briggs, E.R., Fang, B., Everett, L.J., et al. (2015). Genetic Variation Determines PPARgamma Function and Anti-diabetic Drug Response In Vivo. Cell 162, 33-44.

Solar, M., Cardalda, C., Houbracken, I., Martin, M., Maestro, M.A., De Medts, N., Xu, X., Grau, V., Heimberg, H., Bouwens, L., et al. (2009). Pancreatic exocrine duct cells give rise to insulin-producing beta cells during embryogenesis but not after birth. Dev Cell 17, 849-860.

Solloway, M.J., Madjidi, A., Gu, C., Eastham-Anderson, J., Clarke, H.J., Kljavin, N., Zavala-Solorio, J., Kates, L., Friedman, B., Brauer, M., et al. (2015). Glucagon Couples Hepatic Amino Acid Catabolism to mTOR-Dependent Regulation of alpha-Cell Mass. Cell Rep 12, 495-510.

Sosa-Pineda, B., Chowdhury, K., Torres, M., Oliver, G., and Gruss, P. (1997). The Pax4 gene is essential for differentiation of insulin-producing beta cells in the mammalian pancreas. Nature 386, 399-402.

Soyer, J., Flasse, L., Raffelsberger, W., Beucher, A., Orvain, C., Peers, B., Ravassard, P., Vermot, J., Voz, M.L., Mellitzer, G., et al. (2010). Rfx6 is an Ngn3-dependent winged helix transcription factor required for pancreatic islet cell development. Development 137, 203-212.

Spence, J.R., and Wells, J.M. (2007). Translational embryology: using embryonic principles to generate pancreatic endocrine cells from embryonic stem cells. Dev Dyn 236, 3218-3227.

St-Onge, L., Sosa-Pineda, B., Chowdhury, K., Mansouri, A., and Gruss, P. (1997). Pax6 is required for differentiation of glucagon-producing alpha-cells in mouse pancreas. Nature 387, 406-409.

Stagner, J.I., and Samols, E. (1991). Deterioration of Islet Beta-Cell Function After Hemipancreatectomy in Dogs. Diabetes 40, 1472-1479.

Stamateris, R.E., Sharma, R.B., Kong, Y., Ebrahimpour, P., Panday, D., Ranganath, P., Zou, B., Levitt, H., Parambil, N.A., O'Donnell, C.P., et al. (2016). Glucose Induces Mouse b-Cell Proliferation via IRS2, MTOR, and Cyclin D2 but Not the Insulin Receptor. Diabetes 65, 981-995.

Staub A, Sinn L, Behrens OK. Purification and crystallization of the hyperglycemic-glycogenolytic factor (HGF). Science 117: 628, 195

Steiner, D.J., Kim, A., Miller, K., and Hara, M. (2011). Pancreatic islet plasticity: Interspecies comparison of islet architecture and composition. Islets 2, 135-145.

Stratton, I.M., Adler, A.I., Neil, H.A.W., Matthews, D.R., Manley, S.E., Cull, C.A., Hadden, D., Turner, R.C., and Holman, R.R. (2000). Association of glycaemia with macrovascular and microvascular complications of type 2 diabetes (UKPDS 35): prospective observational study. BMJ 321, 405-412.

Suissa, Y., Magenheim, J., Stolovich-Rain, M., Hija, A., Collombat, P., Mansouri, A., Sussel, L., Sosa-Pineda, B., McCracken, K., Wells, J.M., et al. (2013). Gastrin: a distinct fate of neurogenin3 positive progenitor cells in the embryonic pancreas. PLoS One 8, e70397.

Sussel, L., Kalamaras, J., Hartigan, C.R., Meneses, J.J., Pedersen, R.A., Rubenstein, J.L.R., and German, M.S. (1998). Mice lacking the homeodomain transcription factor Nkx2.2 have diabetes due to arrested differentiation of pancreatic β cells. Development 125, 2213-2221.

Swisa, A., Avrahami, D., Eden, N., Zhang, J., Feleke, E., Dahan, T., Cohen-Tayar, Y., Stolovich-Rain, M., Kaestner, K.H., Glaser, B., et al. (2017). PAX6 maintains beta cell identity by repressing genes of alternative islet cell types. J Clin Invest 127, 230-243.

Szklarczyk, D., Morris, J.H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., Santos, A., Doncheva, N.T., Roth, A., Bork, P., et al. (2017). The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res 45, D362-D368.

Szot, G.L., Yadav, M., Lang, J., Kroon, E., Kerr, J., Kadoya, K., Brandon, E.P., Baetge, E.E., Bour-Jordan, H., and Bluestone, J.A. (2015). Tolerance induction and reversal of diabetes in mice transplanted with human embryonic stem cell-derived pancreatic endoderm. Cell Stem Cell 16, 148-157.

Taborsky, G.J., Ahren, B., Mundinger, T.O., Mei, Q., Havel, P.J. (2002). Autonomic mechanism and defects in the glucagon response to insulin-induced hypoglycaemia. Diabetes Nutr Metab. *5*, 318-322.

Taft, R.J., Pheasant, M., and Mattick, J.S. (2007). The relationship between non-protein-coding DNA and eukaryotic complexity. BioEssays *29*, 288–299.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 131, 861-872.

Takahashi, K., and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. Cell 126, 663-676.

Talchai, C., Xuan, S., Lin, H.V., Sussel, L., and Accili, D. (2012). Pancreatic beta cell dedifferentiation as a mechanism of diabetic beta cell failure. Cell 150, 1223-1234.

Tattikota, S.G., Sury, M.D., Rathjen, T., Wessels, H.H., Pandey, A.K., You, X., Becker, C., Chen, W., Selbach, M., and Poy, M.N. (2013). Argonaute2 Regulates the Pancreatic Molecular and Cellular Proteomics 12, 1214-1225.

Taylor, B.L., Liu, F.F., and Sander, M. (2013). Nkx6.1 is essential for maintaining the functional state of pancreatic beta cells. Cell Rep 4, 1262-1275.

Thomson, J.A., Marshall, V.S., and Trojanowski, J.Q. (1998). Neural differentiation of rhesus embryonic stem cells. APMIS 106, 149-157.

Thorel, F., Nepote, V., Avril, I., Kohno, K., Desgraz, R., Chera, S., and Herrera, P.L. (2010). Conversion of adult pancreatic alpha-cells to beta-cells after extreme beta-cell loss. Nature 464, 1149-1154.

Thorens, B., Sarkar, H.K., Kaback, H.R., and Lodish, H.F. (1988). Cloning and Functional Expression in Bacteria of a Novel Glucose Transporter Present in Liver, Intestine, Kidney, and Beta-Pancreatic Islet Cells. Cell 55, 281-290.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 7, 562-578.

Tripathi, V., Ellis, J.D., Shen, Z., Song, D.Y., Pan, O., Watt, A.T., Freier, S.M., Bennett, F., Sharma, A., Bubulya, P.A., Blencowe, B.J., Prasanth, S.G., and Prasanth, K.V. (2010). The Nuclear-Retained Noncoding RNA MALAT1 Regulates Alternative Splicing by Modulating SR Splicing Factor Phosphorylation. Molecular Cell *39*, 925-938.

Ulitsky, I., Shkumatava, A., Jan, C.H., Sive, H., and Bartel, D.P. (2011). Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. Cell 147, 1537-1550.

Unger, R.H., and Cherrington, A.D. (2012). Glucagonocentric restructuring of diabetes: a pathophysiologic and therapeutic makeover. J Clin Invest 122, 4-12.

Unger RH, Eisentraut AM, Mc CM, Keller S, Lanz HC, Madison LL. Glucagon anti- bodies and their use for immunoassay for glucagon. Proc Soc Exp Biol Med 102: 621– 623, 1959.

van Arensbergen, J., Garcia-Hurtado, J., Moran, I., Maestro, M.A., Xu, X., Van de Casteele, M., Skoudy, A.L., Palassini, M., Heimberg, H., and Ferrer, J. (2010). Derepression of Polycomb targets during pancreatic organogenesis allows insulin-producing beta-cells to adopt a neural gene activity program. Genome Res 20, 722-732.

van der Meulen, T., Mawla, A.M., DiGruccio, M.R., Adams, M.W., Nies, V., Dolleman, S., Liu, S., Ackermann, A.M., Caceres, E., Hunter, A.E., et al. (2017). Virgin Beta Cells Persist throughout Life at a Neogenic Niche within Pancreatic Islets. Cell Metab 25, 911-926 e916.

Vance, K.W., Sansom, S.N., Lee, S., Chalei, V., Kong, L., Cooper, S.E., Oliver, P.L., and Ponting, C.P. (2014). The long non-coding RNA Paupar regulates the expression of both local and distal genes. EMBO J 33, 296-311.

Villasenor, A., Chong, D.C., and Cleaver, O. (2008). Biphasic Ngn3 expression in the developing pancreas. Dev Dyn 237, 3270-3279.

Volders, P.J., Helsens, K., Wang, X., Menten, B., Martens, L., Gevaert, K., Vandesompele, J., and Mestdagh, P. (2013). LNCipedia: a database for annotated human lncRNA transcript sequences and structures. Nucleic Acids Res 41, D246-251.

von Mering J. and Minkowski, O. (1890). Diabetes mellitus nach Pankreasextirpation. Archiv. Exp. Pathol. Pharmacol. 26, 371–387.

Wang, H., Bender, A., Wang, P., Karakose, E., Inabnet, W.B., Libutti, S.K., Arnold, A., Lambertini, L., Stang, M., Chen, H., et al. (2017). Insights into beta cell regeneration for diabetes via integration of molecular landscapes in human insulinomas. Nat Commun 8, 767.

Wang, H., Qian, W.J., Mottaz, H.M., Clauss, T.R.W., Anderson, D.J., Moore, R.J., Camp, D.G., Pallavicini, M., Smith, D.J., and Smith, R.D. (2005). Development and Evaluation of a Micro- and Nano-Scale Proteomic Sample Preparation Method. Journal Proteome Research 4, 2397-2403.

Wang, H.W., Breslin, M.B., and Lan, M.S. (2007). Pdx-1 Modulates Histone H4 Acetylation and Insulin Gene Expression in Terminally Differentiated alphaTC-1 Cells. Pancreas 34, 248-253.

Wang, K.C., and Chang, H.Y. (2011). Molecular mechanisms of long noncoding RNAs. Mol Cell 43, 904-914.

Wang, K.C., Yang, Y.W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., Lajoie, B.R., Protacio, A., Flynn, R.A., Gupta, R.A., et al. (2011). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature 472, 120-124.

Wang, L., Park, H.J., Dasari, S., Wang, S., Kocher, J.P., and Li, W. (2013). CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. Nucleic Acids Res 41, e74.

Wang, M., You, J., Bemis, K.G., Tegeler, T.J., and Brown, D.P. (2008). Label-free mass spectrometry-based protein quantification technologies in proteomic analysis. Brief Funct Genomic Proteomic 7, 329-339.

Wang, P., Fiaschi-Taesch, N.M., Vasavada, R.C., Scott, D.K., Garcia-Ocana, A., and Stewart, A.F. (2015). Diabetes mellitus--advances and challenges in human beta-cell proliferation. Nat Rev Endocrinol 11, 201-212.

Wang, Y., Liu, J., Liu, C., Naji, A., and Stoffers, D.A. (2013). MicroRNA-7 Regulates the mTOR Pathway and Proliferation in Adult Pancreatic Beta-Cells. Diabetes 62, 887-895.

Warming, S., Costantino, N., Court, D.L., Jenkins, N.A., and Copeland, N.G. (2005). Simple and highly efficient BAC recombineering using galK selection. Nucleic Acids Res 33, e36.

Watada, H., Scheel, D.W., Leung, J., and German, M.S. (2003). Distinct gene expression programs function in progenitor and mature islet cells. J Biol Chem 278, 17130-17140.

Weedon, M.N., Cebola, I., Patch, A.M., Flanagan, S.E., De Franco, E., Caswell, R., Rodriguez-Segui, S.A., Shaw-Smith, C., Cho, C.H., Allen, H.L., et al. (2014). Recessive mutations in a distal PTF1A enhancer cause isolated pancreatic agenesis. Nat Genet 46, 61-64.

Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., Bernstein, B.E., and Jaenisch, R. (2007). In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. Nature 448, 318-324.

Wessells, N.K., and Cohen, J.H. (1967). Early Pancreas Organogenesis: Morphogenesis, Tissue Interactions, and Mass Effects. Developmental Biology 15, 237-270.

West, J.A., Davis, C.P., Sunwoo, H., Simon, M.D., Sadreyev, R.I., Wang, P.I., Tolstorukov, M.Y., and Kingston, R.E. (2014). The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. Mol Cell 55, 791-802.

Wiebe, P.O., Kormish, J.D., Roper, V.T., Fujitani, Y., Alston, N.I., Zaret, K.S., Wright, C.V., Stein, R.W., and Gannon, M. (2007). Ptf1a binds to and activates area III, a highly conserved region of the Pdx1 promoter that mediates early pancreas-wide Pdx1 expression. Mol Cell Biol 27, 4093-4104.

Wilcox, C.L., Terry, N.A., Walp, E.R., Lee, R.A., and May, C.L. (2013). Pancreatic alpha-cell specific deletion of mouse Arx leads to alpha-cell identity loss. PLoS One 8, e66214.

Wilusz, J.E., Sunwoo, H., and Spector, D.L. (2009). Long noncoding RNAs: functional surprises from the RNA world. Genes Dev 23, 1494-1504.

Wright, J.R. (2002). Almost famous: E. Clark Noble, the common thread in the discovery of insulin and vinblastine. CMAJ 167, 1391-1396.

Xiao, X., Guo, P., Shiota, C., Zhang, T., Coudriet, G.M., Fischbach, S., Prasadan, K., Fusco, J., Ramachandran, S., Witkowski, P., et al. (2018). Endogenous Reprogramming of Alpha Cells into Beta Cells, Induced by Viral Gene Therapy, Reverses Autoimmune Diabetes. Cell Stem Cell 22, 78-90 e74.

Xing, Z., Lin, A., Li, C., Liang, K., Wang, S., Liu, Y., Park, P.K., Qin, L., Wei, Y., Hawke, D.H., et al. (2014). lncRNA directs cooperative epigenetic regulation downstream of chemokine signals. Cell 159, 1110-1125.

Xu, C.R., Cole, P.A., Meyers, D.J., Kormish, J., Dent, S., and Zaret, K.S. (2011). Chromatin "Prepattern" and Histone Modifiers in a Fate Choice for Liver and Pancreas. Science 332, 963-966.

Yang, S.N., and Berggren, P.O. (2006). The role of voltage-gated calcium channels in pancreatic beta-cell physiology and pathophysiology. Endocr Rev 27, 621-676.

Yin, D.D., Zhang, E.B., You, L.H., Wang, N., Wang, L.T., Jin, F.Y., Zhu, Y.N., Cao, L.H., Yuan, Q.X., De, W., et al. (2015). Downregulation of lncRNA TUG1 affects apoptosis and insulin secretion in mouse pancreatic beta cells. Cell Physiol Biochem 35, 1892-1904.

You, L., Wang, N., Yin, D., Wang, L., Jin, F., Zhu, Y., Yuan, Q., and De, W. (2016). Downregulation of Long Noncoding RNA Meg3 Affects Insulin Synthesis and Secretion in Mouse Pancreatic Beta Cells. J Cell Physiol 231, 852-862.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., et al. (2007). Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells. Science 318, 1917-1920.

Zhao, J., Ohsumi, T.K., Kung, J.T., Ogawa, Y., Grau, D.J., Sarma, K., Song, J.J., Kingston, R.E., Borowsky, M., and Lee, J.T. (2010). Genome-wide identification of polycomb-associated RNAs by RIP-seq. Mol Cell 40, 939-953.

Zhao, W., Rasheed, A., Tikkanen, E., Lee, J.J., Butterworth, A.S., Howson, J.M.M., Assimes, T.L., Chowdhury, R., Orho-Melander, M., Damrauer, S., et al. (2017). Identification of new susceptibility loci for type 2 diabetes and shared etiological pathways with coronary heart disease. Nat Genet 49, 1450-1457.

Zhou, Q., and Melton, D.A. (2018). Pancreas regeneration. Nature 557, 351-358.

Zhu, H., Shyh-Chang, N., Segre, A.V., Shinoda, G., Shah, S.P., Einhorn, W.S., Takeuchi, A., Engreitz, J.M., Hagan, J.P., Kharas, M.G., et al. (2011). The Lin28/let-7 axis regulates glucose metabolism. Cell 147, 81-94.

Zhu, Z., Li, Q.V., Lee, K., Rosen, B.P., Gonzalez, F., Soh, C.L., and Huangfu, D. (2016). Genome Editing of Lineage Determinants in Human Pluripotent Stem Cells Reveals Mechanisms of Pancreatic Development and Diabetes. Cell Stem Cell 18, 755-768.

Zisman, A., Peroni, O.D., Abel, E.D., Michael, M.D., Mauvis-Jarvis, F., Lowell, B.B., Wojtaszewski, J.F.P., Hirshman, M.F., Virkamaki, A., Goodyear, L.J., et al. (2000). Targeted disruption of the glucose transporter 4 selectively in muscle causes insulin resistance and glucose intolerance. Nature Medicine 6, 924-928.

Zou, G., Liu, T., Guo, L., Huang, Y., Feng, Y., Huang, Q., and Duan, T. (2016). miR-145 modulates lncRNA-ROR and Sox2 expression to maintain human amniotic epithelial stem cell pluripotency and beta islet-like cell differentiation efficiency. Gene 591, 48-57.

Zuelzer, G. (1908). Uber versuche einer specifischen Fermenttherapie de Diabetes. Zeitschrift fur experimentelle Pathologie und Therapie 5, 307–318.

APPENDIX:

Identification of alternatively spliced genes between alpha and beta cell transcriptomes

All experiments described in the Appendix were conceived and designed by Ruth Singer with guidance from Lori Sussel. Ruth Singer performed all analyses described in this section.

Introduction

Pre-mRNA alternative splicing (AS) is a key post-transcriptional regulatory mechanism that affects gene expression, acting as a major generator of proteomic diversity (Juan-Mateu, 2016). This process regulates the incorporation of alternative sets of exons into mature mRNA molecules, allowing single genes to produce multiple, structurally distinct mRNA and protein isoforms. This tightly regulated process provides cells with the ability to rapidly adapt their transcriptome and proteome in response to intra and extracellular cues. Furthermore, the prevalence and extent of AS correlates with organismal complexity, suggesting that AS plays a key role for the development of complex phenotypic traits during evolution (Kim et al, 2007). AS regulation plays an important role in virtually all biological processes, including cell growth and death, development stage, pluripotency, differentiation, circadian rhythms, response to stimuli and disease. While some studies have suggested a role for AS in the development of T1DM (Juan-Mateu, 2016), genes that undergo AS are only beginning to gain attention in the field of islet biology. Given our findings described in Chapter 2, that *Paupar* interacts with several splicing factors, we decided to analyze alpha and beta cells isolated from mouse and human islets. Through these analyses, we hoped to determine differentially spliced genes that

might help explain our *Paupar* KO alpha cell phenotype and further our understanding of islet cellular identity.

Methods

Replicate Multivariate Analysis of Transcript Splicing (rMATS) is a computational tool to detect differential alternative splicing events from RNA-Seq data (Shen et al., 2014). The statistical model of rMATS calculates the p-value and false discovery rate (FDR) that the difference in the isoform ratio of a gene between two experimental conditions exceeds a statistical threshold. The rMATS algorithm can automatically detect and analyze alternative splicing events corresponding to all major types of alternative splicing patterns, including 5' alternative splice site (5'ASS), 3' alternative splice site (3'ASS), mutually exclusive exons (MXE), skipped exon (SE), and retained intron (RI) (Figure Appendix-1). I analyzed 2 different datasets using rMATS; FACS purified mouse alpha versus beta cells (DiGruccio et al., 2016) and FACS purified human alpha versus beta cells (Ackermann et al., 2016). For these analyses, I used a cutoff p-value of < 0.05 to determine significantly differentially spliced genes between alpha and beta cells.

Results

Differently spliced genes in mouse alpha versus beta cells

To identify genes that are differently spliced between mouse alpha and beta cells, I performed rMATS analyses on RNA-seq data from FACS purified cell populations (DiGruccio et al., 2015). This bioinformatic analyses yielded 2555 alternatively spliced genes (p-value < 0.05) (Figure Appendix-2). Interestingly, the majority of splicing events were skipped exons (1034 genes or 40.47%), followed by retained intron (899 genes or 35.19%), 3' ASS (319 genes 12.49%), 5'

ASS (237 genes or 9.28%), and MXE (66 genes or 2.58%). Only the top 600 alternatively spliced genes (for space considerations) and information about their loci and statistical significance are listed in Table Appendix-1.

Differently spliced genes in human alpha versus beta cells

To identify genes that are differently spliced between mouse alpha and beta cells, I performed rMATS analyses on RNA-seq data from FACS purified cell populations (Ackermann et al., 2016). This bioinformatic analyses yielded 588 alternatively spliced genes (p-value < 0.05) (Figure Appendix-2). Unlike the mouse data, the majority of splicing events were retained introns (236 genes or 40.14%), followed by skipped exons (205 genes or 34.867%), 5' ASS (63 genes or 10.71%), 3' ASS (60 genes 10.20%), and MXE (24 genes or 4.08%). All 588 alternatively spliced genes and information about their loci and statistical significance are listed in Table Appendix-2.

Discussion

Our findings described in this section establish alternative splicing as a mechanism contributing to transcriptional diversity between islet alpha and beta cells. As previously mentioned, we implanted these analyses to further understand the alpha cell dysfunction seen in *Paupar* KO mice. We were surprised and excited to discover that *Pax6* was one of the genes alternatively spliced between mouse alpha and beta cells. Specifically, alpha cells are enriched for *Pax6 5a* isoform, compared to the *Pax6* isoform. While some studies in cell lines have shown that the two splice variants of PAX6 bind distinct DNA motifs (Epstein et al., 1994; Chauhan et al., 2004), the regulatory role of each isoform in islets had yet to be described. Importantly, these analyses for the basis of our mechanistic experiments showing that *Paupar* promotes the alternative

splicing of *Pax6* to the *Pax6 5a* isoform, which is in turn required for the activation of alpha cell genes.

These analyses also uncover surprising differences between human and mouse islet endocrine cells. For example, *Pax6* is significantly alternatively spliced between mouse alpha and beta cells, but not in human alpha versus beta cells. Furthermore, the number of alternative spliced genes between human alpha and beta cells (588) was much lower than the number of genes obtained from mouse alpha and beta cells (2555) (Figure Appendix-2), which could be due to cellular heterogeneity of different human donors leading to larger statistical variation. For my thesis work, downstream analyses of this data was limited to *Pax6*; however, this list of alternatively spliced genes warrant further investigation. Taken together, unveiling the role of AS in specification of cellular identity will increase our understanding of islet function and dysfunction and may lead to new strategies for disease prevention and therapy.

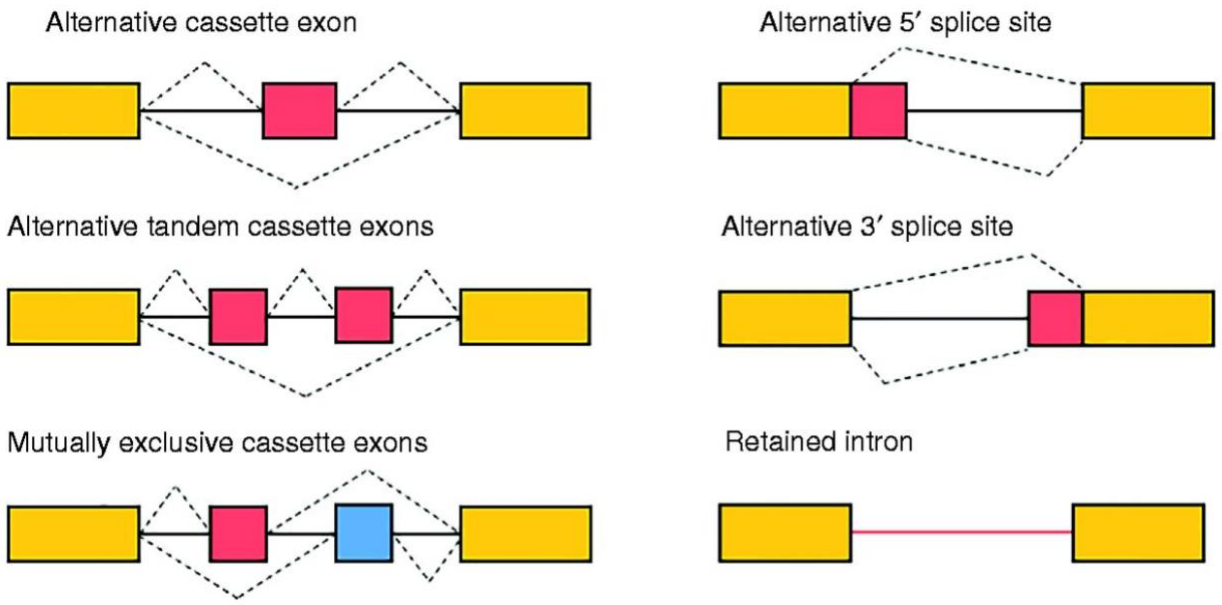


Figure Appendix-1. Regulation of alternative splicing (AS). Major types of AS events. Exons are represented by boxes and introns by solid lines; constitutive exons are shown in yellow while alternative exons are shown in red or blue; dashed lines represent different splicing events. Adapted from Juan-Mateu, 2016.

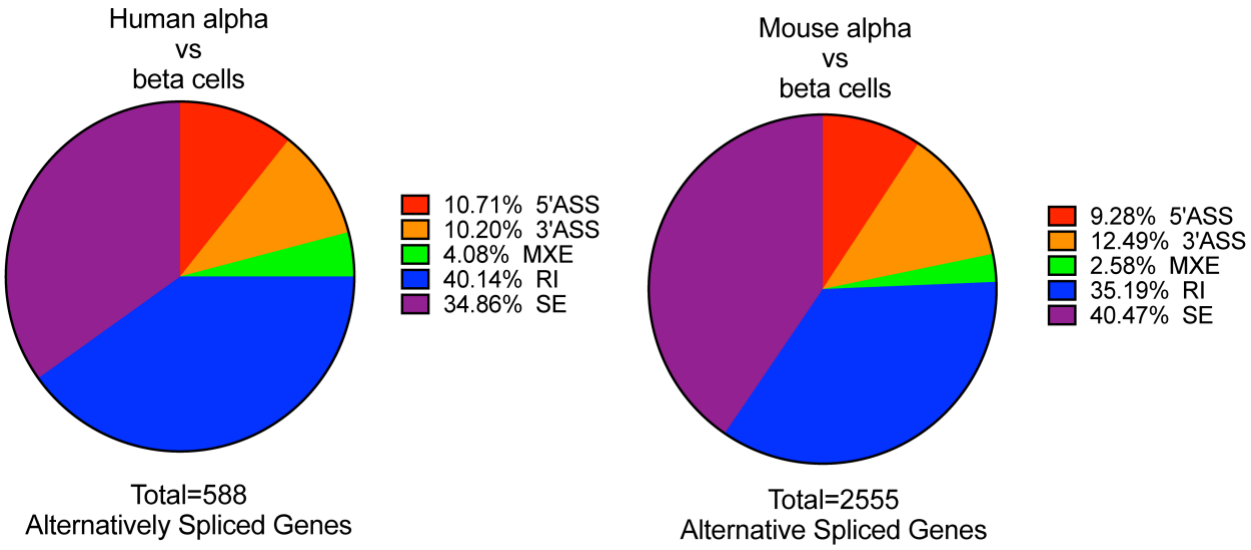


Figure Appendix-2. Breakdown of significantly spliced genes in mouse and human alpha and beta cells by type of splicing event. 5' alternative splice site (5'ASS), 3' alternative splice site (3'ASS), mutually exclusive exons (MXE), skipped exon (SE), and retained intron (RI).

Table Appendix-1. List of genes significantly alternatively spliced between mouse alpha and beta cells. Only the top 600 genes are included below.

Splicing Event	Gene	Chr	Strand	Exon_Start	Exon_End	PValue	FDR
RI	Pafah1b1	chr11	-	74673949	74679614	0.000000	0.000000
RI	Tmem57	chr4	-	134802758	134806691	0.000000	0.000000
RI	Anapc5	chr5	-	122799338	122800593	0.000000	0.000000
RI	Sec11c	chr18	+	65812664	65814957	0.000000	0.000000
RI	Eif4a2	chr16	+	23110579	23111263	0.000000	0.000000
RI	Eif4a2	chr16	+	23111125	23111592	0.000000	0.000000
RI	Rbm3	chrX	-	8143847	8144679	0.000000	0.000000
RI	Enpp5	chr17	+	44082743	44086567	0.000000	0.000000
RI	Maged1	chrX	-	94537675	94538065	0.000000	0.000000
RI	Atp6v0a1	chr11	+	101027266	101029258	0.000000	0.000000
SE	Bicdl1	chr5	-	115651155	115651264	0.000000	0.000000
SE	Rps24	chr14	+	24495429	24495449	0.000000	0.000000
SE	Rbm3	chrX	-	8144174	8144445	0.000000	0.000000
SE	Ptprf	chr4	-	118218011	118218044	0.000000	0.000000
SE	Gnas	chr2	+	174336746	174336791	0.000000	0.000000
RI	Mt1	chr8	+	94179822	94180327	0.000000	0.000000
SE	Srsf1	chr11	+	88049411	88049598	0.000000	0.000000
SE	Clta	chr4	+	44031399	44031435	0.000000	0.000000
RI	Rab3gap2	chr1	+	185275911	185277221	0.000000	0.000000
5' ASS	Gnas	chr2	+	174328107	174330434	0.000000	0.000000
RI	Copa	chr1	+	172091163	172092354	0.000000	0.000000
RI	Bcas2	chr3	+	103175618	103177419	0.000000	0.000000
RI	Celf3	chr3	+	94486793	94487284	0.000000	0.000000
RI	Sec31a	chr5	-	100363376	100363947	0.000000	0.000000
RI	Abcc8	chr7	-	46107237	46108214	0.000000	0.000000
RI	Mapkapk3	chr9	-	107262422	107263764	0.000000	0.000000
SE	Morf4l2	chrX	-	136740945	136741119	0.000000	0.000000
RI	Pafah1b1	chr11	-	74684405	74686064	0.000000	0.000000
MXE	Slc7a2	chr8	+	40908055	40908195	0.000000	0.000000
RI	Mpp3	chr11	-	102023701	102025154	0.000000	0.000000
RI	Dnaja1	chr4	+	40726104	40728183	0.000000	0.000000
RI	Nrip2	chr6	+	128404934	128406572	0.000000	0.000000
RI	Eif4a2	chr16	+	23109984	23110343	0.000000	0.000000
SE	Nptn	chr9	+	58618662	58619010	0.000000	0.000000
RI	St7	chr6	+	17850356	17852412	0.000000	0.000000
RI	Arglu1	chr8	-	8681383	8683973	0.000000	0.000000
RI	Dcaf6	chr1	-	165338187	165340023	0.000000	0.000000
RI	Napb	chr2	-	148707158	148709433	0.000000	0.000000

SE	Sorbs2	chr8	+	45794977	45796558	0.000000	0.000000
3' ASS	Banp	chr8	+	122007090	122007162	0.000000	0.000000
RI	Tmem87b	chr2	+	128845460	128846799	0.000000	0.000000
RI	43350	chr9	+	25303716	25306265	0.000000	0.000000
SE	Peg3	chr7	-	6712949	6713012	0.000000	0.000000
RI	Thoc7	chr14	-	13953410	13954656	0.000000	0.000000
RI	Klhdc10	chr6	+	30447940	30449630	0.000000	0.000000
SE	Syne2	chr12	+	76100027	76100096	0.000000	0.000000
SE	Stox1	chr10	-	62724262	62724432	0.000000	0.000000
RI	Ugdh	chr5	-	65420248	65422782	0.000000	0.000000
SE	Arglu1	chr8	-	8681840	8681948	0.000000	0.000000
RI	Nupl1	chr14	-	60240773	60242657	0.000000	0.000000
SE	Uqcr10	chr11	-	4703903	4703989	0.000000	0.000000
5' ASS	Piga	chrX	+	164422583	164423672	0.000000	0.000000
SE	Clta	chr4	+	44030214	44030268	0.000000	0.000000
RI	Hypk	chr2	+	121457645	121458043	0.000000	0.000000
RI	Snx5	chr2	-	144257099	144257993	0.000000	0.000000
3' ASS	Ogg1	chr6	+	113329145	113329408	0.000000	0.000000
RI	Eif5b	chr1	+	38050186	38051284	0.000000	0.000000
SE	Khk	chr5	+	30927021	30927156	0.000000	0.000000
RI	Becn1	chr11	-	101294000	101295164	0.000000	0.000000
RI	43352	chr11	+	117356612	117359470	0.000000	0.000000
RI	Cdk4	chr10	+	127065154	127066003	0.000000	0.000000
RI	Srprb	chr9	-	103197537	103198866	0.000000	0.000000
RI	Becn1	chr11	-	101289782	101290499	0.000000	0.000000
RI	Gdpd3	chr7	+	126770987	126771213	0.000000	0.000000
RI	Eif4a1	chr11	-	69667694	69668111	0.000000	0.000000
MXE	Khk	chr5	+	30926687	30926822	0.000000	0.000000
RI	Krt7	chr15	+	101413968	101416281	0.000000	0.000000
RI	Fam32a	chr8	+	72219940	72221981	0.000000	0.000000
RI	Abcc8	chr7	-	46105808	46106688	0.000000	0.000000
RI	Adipor1	chr1	+	134423957	134424922	0.000000	0.000000
RI	Txnrd1	chr10	+	82881722	82884006	0.000000	0.000000
RI	Eif4a1	chr11	-	69668195	69668548	0.000000	0.000000
RI	Pmpcb	chr5	+	21743378	21746090	0.000000	0.000000
RI	Clcn3	chr8	-	60922789	60927418	0.000000	0.000000
MXE	Slc39a14	chr14	-	70313600	70313770	0.000000	0.000000
SE	Ctage5	chr12	+	59144698	59144770	0.000000	0.000000
RI	Atg4b	chr1	+	93775627	93778364	0.000000	0.000000
RI	Anxa11	chr14	+	25872511	25874031	0.000000	0.000000
SE	Rps19	chr7	+	24886106	24886187	0.000000	0.000000

SE	Shc1	chr3	+	89424189	89424243	0.000000	0.000000
5' ASS	Sri	chr5	+	8062383	8063000	0.000000	0.000000
3' ASS	Tpm1	chr9	-	67022592	67024565	0.000000	0.000001
RI	Mbtps1	chr8	-	119538843	119541694	0.000000	0.000000
RI	M6pr	chr6	+	122313258	122315166	0.000000	0.000000
RI	Eif4a2	chr16	+	23108587	23108851	0.000000	0.000000
RI	Ppib	chr9	+	66062988	66065671	0.000000	0.000000
RI	Ipo7	chr7	+	110044617	110046251	0.000000	0.000000
SE	43167	chr6	+	116401336	116402182	0.000000	0.000000
5' ASS	Nop56	chr2	+	130277031	130277314	0.000000	0.000000
SE	Usp54	chr14	-	20581456	20581627	0.000000	0.000000
SE	Rab12	chr17	-	66504988	66505178	0.000000	0.000000
3' ASS	Arid5a	chr1	+	36317203	36317662	0.000000	0.000001
SE	Epb41l3	chr17	+	69283884	69284007	0.000000	0.000001
MXE	Slc25a3	chr10	-	91121921	91122043	0.000000	0.000000
RI	Gdi1	chrX	+	74309075	74310134	0.000000	0.000000
RI	Lmbrd1	chr1	+	24743882	24744358	0.000000	0.000000
RI	Hmgn3	chr9	-	83109947	83111106	0.000000	0.000000
3' ASS	Nsun3	chr16	-	62776287	62776631	0.000000	0.000001
SE	Ptprf	chr4	-	118227818	118227845	0.000000	0.000001
SE	2310035C23Rik	chr1	+	105741012	105741095	0.000000	0.000001
RI	Rnf167	chr11	+	70649674	70650017	0.000000	0.000000
SE	Lrch3	chr16	+	32995805	32995913	0.000000	0.000001
RI	Anapc2	chr2	+	25272914	25273932	0.000000	0.000000
RI	Ptprd	chr4	-	76085460	76086599	0.000000	0.000000
RI	Ddx5	chr11	-	106782468	106784018	0.000000	0.000000
3' ASS	Eef1b2	chr1	+	63178022	63178565	0.000000	0.000002
RI	Eef2	chr10	+	81181227	81181625	0.000000	0.000000
SE	Atp6v0a1	chr11	+	101049458	101049476	0.000000	0.000002
SE	Oas1b	chr5	+	120813210	120813362	0.000000	0.000002
RI	Srsf5	chr12	+	80947795	80949168	0.000000	0.000000
SE	Psme4	chr11	+	30866035	30866279	0.000000	0.000002
SE	R3hdm2	chr10	+	127462187	127462283	0.000000	0.000002
RI	Mrps10	chr17	+	47372575	47375109	0.000000	0.000000
SE	Ppfibp1	chr6	+	146999661	146999694	0.000000	0.000003
RI	Sharpin	chr15	-	76348085	76348435	0.000000	0.000001
SE	Hmgn3	chr9	-	83110787	83110874	0.000000	0.000004
SE	Gm26881	chr4	+	43052944	43053051	0.000000	0.000004
RI	Srsf6	chr2	+	162931993	162933550	0.000000	0.000001
SE	Khk	chr5	+	30926687	30926822	0.000000	0.000004
RI	Adam8	chr7	-	139987178	139987737	0.000000	0.000001

SE	Baz2b	chr2	-	59983781	59983970	0.000000	0.000005
RI	Bcas2	chr3	+	103174339	103177419	0.000000	0.000001
SE	Spata13	chr14	+	60636367	60636646	0.000000	0.000006
SE	Inpp4a	chr1	+	37383338	37383371	0.000000	0.000006
RI	Utp4	chr8	+	106917581	106918805	0.000000	0.000001
RI	Hps1	chr19	-	42756257	42757906	0.000000	0.000001
RI	Ufd1	chr16	+	18826254	18827106	0.000000	0.000001
RI	Trappc11	chr8	-	47513282	47513808	0.000000	0.000001
RI	0610037L13Rik	chr4	+	107894872	107897118	0.000000	0.000001
RI	Taz	chrX	+	74288177	74288513	0.000000	0.000001
SE	Adgrl1	chr8	+	83924205	83924355	0.000000	0.000008
RI	Pak1ip1	chr13	+	41009447	41011282	0.000000	0.000001
RI	Copg1	chr6	+	87892632	87893953	0.000000	0.000001
5' ASS	Capn7	chr14	+	31360124	31361245	0.000000	0.000012
RI	Dnajc8	chr4	+	132544055	132544751	0.000000	0.000001
RI	Ptpa	chr2	+	30438228	30439764	0.000000	0.000001
RI	Bub3	chr7	+	131561536	131563485	0.000000	0.000001
SE	Etl4	chr2	+	20759546	20759651	0.000000	0.000011
RI	Ttc39a	chr4	+	109422445	109422942	0.000000	0.000002
5' ASS	Thoc7	chr14	-	13954538	13954656	0.000000	0.000012
SE	Snrnp70	chr7	-	45381824	45383184	0.000000	0.000011
RI	Rnpepl1	chr1	+	92916710	92917746	0.000000	0.000002
RI	Cnbp	chr6	-	87845121	87845793	0.000000	0.000002
5' ASS	Gtf2a2	chr9	+	70016818	70017191	0.000000	0.000012
SE	Srsf6	chr2	+	162932814	162933082	0.000000	0.000012
RI	Map1lc3b	chr8	+	121593489	121596092	0.000000	0.000002
5' ASS	Firre	chrX	-	50580580	50580949	0.000000	0.000012
SE	Arhgap21	chr2	-	20853957	20854232	0.000000	0.000013
RI	Nabp1	chr1	-	51477477	51478399	0.000000	0.000002
RI	Apbb1	chr7	-	105565532	105565951	0.000000	0.000002
SE	R3hdm2	chr10	+	127462205	127462283	0.000000	0.000014
RI	N4bp1	chr8	-	86846814	86848554	0.000000	0.000002
RI	Nrip2	chr6	+	128407685	128408932	0.000000	0.000002
RI	Diablo	chr5	-	123517736	123518065	0.000000	0.000002
5' ASS	Alkbh8	chr9	+	3335477	3335748	0.000000	0.000015
5' ASS	2900076A07Rik	chr7	+	81523570	81523723	0.000000	0.000015
RI	Aebp2	chr6	+	140650755	140652826	0.000000	0.000003
RI	Txndc9	chr1	-	37993991	37995841	0.000000	0.000003
RI	Tmem214	chr5	+	30871419	30871855	0.000000	0.000003
SE	Spire1	chr18	-	67512840	67512879	0.000000	0.000021
RI	Lss	chr10	+	76544464	76545501	0.000000	0.000003

RI	Stk25	chr1	-	93626050	93627133	0.000000	0.000003
3' ASS	Grin1	chr2	-	25291180	25292484	0.000000	0.000046
RI	Usp7	chr16	-	8697892	8698586	0.000000	0.000004
SE	Ap3d1	chr10	-	80710088	80710292	0.000000	0.000025
RI	Mocs2	chr13	+	114825210	114826281	0.000000	0.000004
RI	Rabggtb	chr3	-	153910260	153910888	0.000000	0.000004
RI	Timm10b	chr7	+	105640761	105641228	0.000000	0.000004
3' ASS	Celf2	chr2	-	6539704	6546779	0.000000	0.000055
5' ASS	Nabp1	chr1	-	51477721	51478399	0.000000	0.000025
RI	Slc30a9	chr5	+	67342093	67344569	0.000000	0.000005
RI	Gas5	chr1	+	161035982	161036244	0.000000	0.000005
RI	Sppl2b	chr10	+	80866499	80868708	0.000000	0.000005
RI	Nub1	chr5	+	24694699	24695590	0.000000	0.000006
SE	Gnas	chr2	+	174328107	174330434	0.000000	0.000043
RI	Tmbim1	chr1	-	74291430	74291734	0.000000	0.000006
RI	Arhgef2	chr3	+	88632912	88633367	0.000000	0.000006
RI	Strip1	chr3	-	107615760	107616608	0.000000	0.000006
SE	Orc3	chr4	-	34598982	34599140	0.000000	0.000047
MXE	Slc25a3	chr10	-	91121921	91122043	0.000000	0.000026
3' ASS	Srsf5	chr12	+	80947115	80947381	0.000000	0.000077
RI	Habp4	chr13	+	64182106	64184758	0.000000	0.000007
RI	Rnpepl1	chr1	+	92917146	92917746	0.000000	0.000007
RI	Rundc3a	chr11	+	102399614	102400045	0.000000	0.000007
3' ASS	Snrnp70	chr7	-	45380716	45383184	0.000000	0.000078
RI	Sdccag3	chr2	-	26385497	26386063	0.000000	0.000008
RI	Rpl10	chrX	+	74271659	74271907	0.000000	0.000008
RI	Zfp940	chr7	-	29846468	29846960	0.000000	0.000008
RI	Mapre3	chr5	+	30862497	30863405	0.000000	0.000008
SE	Armc10	chr5	+	21647307	21647448	0.000000	0.000060
SE	Magi1	chr6	-	93704276	93704360	0.000000	0.000060
SE	Ahi1	chr10	+	21070287	21070543	0.000000	0.000067
RI	Cnn3	chr3	+	121449948	121450358	0.000000	0.000009
RI	Hnrnph1	chr11	+	50379467	50379964	0.000000	0.000009
SE	1810014B01Rik	chr10	+	86688533	86689138	0.000000	0.000067
RI	Med24	chr11	-	98716413	98718142	0.000000	0.000009
RI	Stard7	chr2	+	127284094	127284581	0.000000	0.000010
SE	Srsf3	chr17	+	29039453	29039909	0.000000	0.000068
RI	Gas5	chr1	+	161036207	161036579	0.000001	0.000010
SE	Atp13a3	chr16	-	30321805	30321895	0.000001	0.000068
SE	Srsf6	chr2	+	162932814	162932979	0.000001	0.000068
SE	Nf2	chr11	-	4780577	4780622	0.000001	0.000068

RI	Dhx30	chr9	-	110115035	110115582	0.000001	0.000010
RI	Neu1	chr17	+	34932565	34934185	0.000001	0.000010
SE	Srrm2	chr17	+	23811816	23812077	0.000001	0.000070
SE	Wbp1	chr6	-	83120771	83120876	0.000001	0.000070
5' ASS	Kat2a	chr11	-	100710249	100710636	0.000001	0.000062
RI	Gdi1	chrX	+	74307671	74308261	0.000001	0.000012
SE	Plekha1	chr7	+	130904318	130904557	0.000001	0.000081
3' ASS	Crlf2	chr5	-	109555506	109555660	0.000001	0.000139
RI	Hace1	chr10	+	45700953	45709994	0.000001	0.000013
RI	Lrrc56	chr7	+	141196981	141198361	0.000001	0.000014
RI	Cfap20	chr8	-	95422782	95424696	0.000001	0.000015
3' ASS	Tnfsfm13	chr11	-	69684217	69684336	0.000001	0.000148
RI	Srrm1	chr4	-	135344498	135346857	0.000001	0.000015
SE	Elavl4	chr4	-	110255418	110255505	0.000001	0.000106
RI	Zswim8	chr14	+	20719379	20720036	0.000001	0.000016
SE	Itgb3bp	chr4	-	99802096	99802232	0.000001	0.000106
SE	Pex5l	chr3	-	33024695	33024800	0.000001	0.000106
RI	Ppp1cc	chr5	+	122172731	122173280	0.000001	0.000017
SE	Wbp1	chr6	-	83120771	83120873	0.000001	0.000108
RI	Pdhb	chr14	-	8171422	8172817	0.000001	0.000018
RI	Slc31a1	chr4	+	62387866	62391769	0.000001	0.000019
SE	Gba	chr3	+	89203191	89203360	0.000001	0.000120
RI	Tpr	chr1	+	150447511	150449238	0.000001	0.000019
5' ASS	Celf3	chr3	+	94487134	94487388	0.000001	0.000106
RI	Mroh1	chr15	+	76421139	76423835	0.000001	0.000019
RI	Cul5	chr9	-	53632272	53633839	0.000001	0.000019
RI	R3hdm1	chr1	+	128181985	128184654	0.000001	0.000019
RI	Stk16	chr1	+	75212599	75213006	0.000001	0.000020
SE	Wnk1	chr6	-	119953770	119954232	0.000001	0.000130
SE	Mcfd2	chr17	-	87258749	87258881	0.000001	0.000131
SE	Nnat	chr2	+	157560447	157560535	0.000001	0.000137
RI	Thoc7	chr14	-	13953410	13954656	0.000001	0.000022
3' ASS	Psap	chr10	+	60299091	60299229	0.000001	0.000223
SE	Arid5a	chr1	+	36317203	36317324	0.000001	0.000146
5' ASS	Taz	chrX	+	74288450	74289019	0.000001	0.000133
SE	Atp11a	chr8	+	12861639	12861724	0.000001	0.000156
RI	Cops8	chr1	+	90604374	90605512	0.000002	0.000026
SE	Npr2	chr4	+	43643334	43643409	0.000002	0.000165
RI	Orc3	chr4	-	34597307	34598756	0.000002	0.000027
5' ASS	Ccdc84	chr9	-	44413442	44413533	0.000002	0.000138
RI	Xab2	chr8	-	3613158	3613499	0.000002	0.000027

RI	Trim2	chr3	-	84190512	84192261	0.000002	0.000027
RI	Cacfd1	chr2	+	27018887	27021089	0.000002	0.000028
RI	Xpc	chr6	-	91496087	91498190	0.000002	0.000028
RI	Ddx23	chr15	-	98650143	98650818	0.000002	0.000028
SE	Tm2d3	chr7	+	65693865	65693943	0.000002	0.000185
RI	Gas5	chr1	+	161038224	161038539	0.000002	0.000029
RI	Rps18	chr17	-	33952220	33952591	0.000002	0.000029
RI	Nabp1	chr1	-	51477477	51478399	0.000002	0.000029
RI	Vac14	chr8	+	110712747	110715577	0.000002	0.000030
SE	Jade3	chrX	+	20487297	20487422	0.000002	0.000200
RI	Timm44	chr8	-	4265225	4266680	0.000002	0.000031
3' ASS	Scnn1b	chr7	+	121917349	121917575	0.000002	0.000341
SE	Mynn	chr3	+	30612468	30612552	0.000002	0.000218
5' ASS	Nabp1	chr1	-	51477721	51478399	0.000002	0.000176
RI	Smurf1	chr5	-	144880624	144881819	0.000002	0.000035
RI	Cers4	chr8	+	4519462	4520682	0.000002	0.000035
3' ASS	Mrpl4	chr9	+	21007363	21007540	0.000002	0.000351
SE	Scmh1	chr4	+	120462012	120462155	0.000002	0.000236
SE	Zfp322a	chr13	-	23362579	23362654	0.000002	0.000236
RI	Anxa11	chr14	+	25874209	25874784	0.000002	0.000038
RI	Fbxw7	chr3	+	84967439	84969301	0.000003	0.000039
SE	Tmem234	chr4	+	129601406	129601625	0.000003	0.000252
SE	Sipa1l1	chr12	+	82422660	82423095	0.000003	0.000252
RI	Flt3l	chr7	-	45132227	45134026	0.000003	0.000040
RI	Fancg	chr4	-	43006983	43007403	0.000003	0.000041
RI	Celf2	chr2	-	6539704	6547213	0.000003	0.000041
SE	Fdft1	chr14	-	63159237	63159359	0.000003	0.000261
SE	Gas5	chr1	+	161037123	161037164	0.000003	0.000264
SE	Dis3l2	chr1	+	86812511	86812589	0.000003	0.000270
RI	Ist1	chr8	-	109675375	109676850	0.000003	0.000044
SE	Fbxo34	chr14	+	47500793	47500870	0.000003	0.000277
5' ASS	Nabp1	chr1	-	51477824	51478399	0.000003	0.000255
RI	Mau2	chr8	-	70028618	70029235	0.000004	0.000054
RI	Zc3h14	chr12	+	98785317	98785741	0.000004	0.000055
3' ASS	Rnf216	chr5	-	143098378	143098505	0.000004	0.000544
5' ASS	Pts	chr9	-	50525152	50525284	0.000004	0.000292
SE	Trrap	chr5	+	144810150	144810204	0.000004	0.000377
RI	Idh3g	chrX	-	73780573	73781030	0.000004	0.000063
RI	Usp5	chr6	-	124816983	124817471	0.000005	0.000067
RI	Lcmt1	chr7	+	123421537	123422798	0.000005	0.000071
RI	Ddx54	chr5	+	120624537	120625844	0.000005	0.000072

SE	Eno2	chr6	-	124766530	124766664	0.000005	0.000451
RI	Eno2	chr6	-	124762652	124763206	0.000005	0.000074
3' ASS	Mrpl4	chr9	+	21007327	21007540	0.000005	0.000671
SE	Khk	chr5	+	30926687	30926822	0.000005	0.000451
SE	Scamp5	chr9	-	57447076	57447233	0.000005	0.000467
RI	Mms19	chr19	-	41949454	41949837	0.000006	0.000080
RI	Mta2	chr19	+	8946266	8946750	0.000006	0.000081
SE	Macf1	chr4	-	123354478	123354580	0.000006	0.000496
RI	Cdk5	chr5	-	24422331	24422558	0.000006	0.000085
5' ASS	Sat1	chrX	-	155215119	155215684	0.000006	0.000424
SE	Pcmtd1	chr1	+	7155079	7155271	0.000006	0.000532
RI	Dnajc2	chr5	-	21761540	21763515	0.000006	0.000090
SE	Hnrnpdl	chr5	-	100035634	100035739	0.000007	0.000549
RI	Camk2g	chr14	-	20764860	20765479	0.000007	0.000093
SE	Nisch	chr14	-	31182498	31182640	0.000007	0.000555
RI	Slc35b4	chr6	-	34163588	34166423	0.000007	0.000093
RI	Rhbdf2	chr11	-	116605242	116606232	0.000007	0.000093
SE	Khk	chr5	+	30927021	30927156	0.000007	0.000556
RI	Usp11	chrX	+	20716746	20717804	0.000007	0.000093
RI	Wdr75	chr1	+	45817262	45818264	0.000007	0.000097
SE	Senp7	chr16	+	56187469	56187570	0.000007	0.000583
5' ASS	Nabp1	chr1	-	51478300	51478399	0.000007	0.000472
RI	Atl2	chr17	-	79850886	79852956	0.000007	0.000099
RI	Mbtps1	chr8	-	119515577	119518279	0.000008	0.000103
5' ASS	4833420G17Rik	chr13	+	119473832	119474194	0.000008	0.000479
RI	Nol11	chr11	-	107180041	107181101	0.000008	0.000105
SE	Atg13	chr2	-	91679919	91679997	0.000008	0.000631
RI	Ppfibp1	chr6	+	147010862	147012518	0.000008	0.000105
SE	Rbm39	chr2	-	156177751	156177906	0.000008	0.000631
5' ASS	Zfp446	chr7	+	12979884	12980993	0.000008	0.000494
RI	Psmd3	chr11	+	98690462	98691019	0.000008	0.000112
MXE	Wnk2	chr13	-	49044373	49044418	0.000008	0.000613
SE	Hps5	chr7	-	46775153	46775284	0.000009	0.000677
SE	Tnk2	chr16	+	32680899	32680995	0.000009	0.000688
5' ASS	Slc25a37	chr14	-	69248882	69249622	0.000009	0.000505
RI	Sf3b1	chr1	-	55014482	55016490	0.000009	0.000125
5' ASS	Pabpc4	chr4	+	123295201	123295321	0.000010	0.000519
3' ASS	Rtel1	chr2	+	181355309	181355552	0.000010	0.001187
RI	Slc25a37	chr14	-	69247580	69249622	0.000010	0.000126
RI	Ddx46	chr13	+	55659019	55660799	0.000010	0.000126
SE	Prr3	chr17	-	35978710	35978781	0.000010	0.000756

SE	Mtmr14	chr6	+	113277830	113277945	0.000010	0.000756
SE	Sft2d2	chr1	-	165186721	165186888	0.000010	0.000762
SE	Rap1gap	chr4	+	137723745	137723823	0.000010	0.000762
RI	Tmem214	chr5	+	30874850	30875886	0.000010	0.000133
RI	Nabp1	chr1	-	51469819	51471411	0.000010	0.000133
RI	Vps11	chr9	-	44353836	44354291	0.000011	0.000135
SE	Josd2	chr7	+	44469982	44470172	0.000011	0.000779
RI	Ipo7	chr7	+	110048944	110049693	0.000011	0.000135
RI	Becn1	chr11	-	101295533	101296316	0.000011	0.000135
SE	Dnajc24	chr2	-	105980981	105981120	0.000011	0.000788
SE	Rbm5	chr9	-	107755852	107755913	0.000011	0.000788
5' ASS	Rnf141	chr7	-	110837076	110837270	0.000011	0.000567
3' ASS	Mtmr7	chr8	-	40554270	40554601	0.000011	0.001280
SE	Mdm1	chr10	+	118156833	118156863	0.000012	0.000832
5' ASS	Angel2	chr1	+	190940873	190941458	0.000012	0.000582
SE	Hnrnpd	chr5	-	99962097	99962204	0.000012	0.000844
3' ASS	Mus81	chr19	-	5483634	5483763	0.000012	0.001335
RI	Taz	chrX	+	74288177	74288513	0.000012	0.000157
3' ASS	Zfp68	chr5	-	138616469	138616614	0.000013	0.001349
RI	Hsd3b7	chr7	+	127801074	127801616	0.000013	0.000162
SE	Pabpc4	chr4	+	123295749	123295836	0.000014	0.000955
3' ASS	Gigyf2	chr1	+	87407534	87407723	0.000014	0.001354
3' ASS	1500009L16Rik	chr10	+	83759592	83759630	0.000014	0.001354
RI	Msmo1	chr8	-	64722463	64723736	0.000014	0.000176
5' ASS	Usp35	chr7	-	97315079	97315178	0.000014	0.000685
SE	Ggps1	chr13	-	14058923	14059211	0.000014	0.001006
5' ASS	Gtf2a2	chr9	+	70016818	70017191	0.000015	0.000687
SE	Mgat1	chr11	+	49251549	49251674	0.000015	0.001024
SE	Rbm6	chr9	-	107861622	107861729	0.000015	0.001025
SE	Tm2d3	chr7	+	65693850	65693943	0.000016	0.001052
SE	Mtmr14	chr6	+	113277789	113277945	0.000016	0.001063
RI	Slc50a1	chr3	-	89269826	89270139	0.000016	0.000197
RI	1810022K09Rik	chr3	-	14606288	14607293	0.000016	0.000197
RI	Rab4b	chr7	-	27175778	27176238	0.000017	0.000202
RI	Pigq	chr17	-	25931673	25932231	0.000017	0.000203
3' ASS	Dclk2	chr3	-	86816357	86816433	0.000017	0.001565
SE	Anapc16	chr10	-	59996447	59996614	0.000017	0.001159
5' ASS	Trerf1	chr17	+	47319351	47319678	0.000018	0.000789
RI	Ilf2	chr3	+	90486850	90487097	0.000018	0.000218
SE	Ppp2r1b	chr9	+	50873196	50873308	0.000018	0.001193
SE	Mcfd2	chr17	-	87264140	87264365	0.000018	0.001196

RI	Ttc12	chr9	-	49471861	49472543	0.000019	0.000224
5' ASS	Tmem234	chr4	+	129600982	129601243	0.000019	0.000843
RI	Ostf1	chr19	-	18584641	18586625	0.000020	0.000236
3' ASS	Uap1	chr1	-	170147995	170148046	0.000020	0.001755
RI	Rccd1	chr7	-	80320301	80321181	0.000020	0.000242
SE	Entpd5	chr12	-	84399339	84399412	0.000021	0.001336
SE	Atp9b	chr18	-	80736811	80736844	0.000021	0.001336
RI	Ufl1	chr4	-	25269027	25270618	0.000021	0.000248
RI	Dusp12	chr1	-	170880129	170881041	0.000021	0.000248
SE	Ctbp2	chr7	-	133025712	133025872	0.000021	0.001340
SE	Nrf1	chr6	+	30140215	30140271	0.000021	0.001340
SE	43354	chr5	+	93173469	93173537	0.000021	0.001345
5' ASS	Ankrd16	chr2	+	11781476	11781533	0.000022	0.000896
RI	Ipo8	chr6	-	148770682	148775093	0.000022	0.000253
RI	Eif4a1	chr11	-	69668195	69668753	0.000022	0.000253
RI	Pex14	chr4	-	148967523	148970635	0.000022	0.000255
5' ASS	Nfatc1	chr18	-	80635401	80636076	0.000022	0.000896
SE	Mcf2l	chr8	+	13013914	13014007	0.000022	0.001392
RI	Myl6	chr10	-	128490860	128492136	0.000023	0.000264
RI	Mat2a	chr6	-	72432798	72435114	0.000023	0.000265
3' ASS	Trp53i13	chr11	-	77509075	77509367	0.000023	0.001982
RI	Zranb2	chr3	+	157541699	157543244	0.000024	0.000278
RI	Leng8	chr7	+	4144052	4144813	0.000024	0.000278
SE	Glis3	chr19	-	28540200	28540408	0.000026	0.001579
RI	Cers4	chr8	+	4519462	4520308	0.000026	0.000295
RI	Npepl1	chr2	+	174120543	174121124	0.000026	0.000296
RI	Dcaf11	chr14	+	55565097	55565541	0.000026	0.000297
RI	Trafd1	chr5	-	121373137	121374067	0.000026	0.000297
RI	Napa	chr7	+	16113349	16113560	0.000027	0.000299
SE	Os9	chr10	-	127096914	127097079	0.000027	0.001673
3' ASS	Pts	chr9	-	50524722	50524964	0.000027	0.002205
3' ASS	1300002E11Rik	chr16	+	21800714	21800836	0.000028	0.002205
RI	Usp19	chr9	+	108493560	108494238	0.000028	0.000308
MXE	Adam22	chr5	-	8095072	8095180	0.000028	0.001745
SE	Samd4	chr14	+	47052849	47053113	0.000030	0.001781
SE	Ctage5	chr12	+	59146335	59146362	0.000030	0.001781
SE	Slmap	chr14	-	26416202	26416292	0.000030	0.001781
RI	Tcirg1	chr19	-	3901890	3902556	0.000030	0.000330
SE	Srrt	chr5	-	137301317	137301647	0.000031	0.001802
RI	Stau1	chr2	-	166951270	166952472	0.000031	0.000338
SE	Dctn4	chr18	+	60545356	60545377	0.000031	0.001823

RI	Faah	chr4	-	115999498	116000088	0.000031	0.000342
3' ASS	Vps50	chr6	+	3601367	3603315	0.000031	0.002413
RI	Mtfr1l	chr4	-	134530274	134530789	0.000033	0.000361
RI	Ncaph2	chr15	+	89370394	89370675	0.000033	0.000361
RI	Cap1	chr4	-	122862622	122863085	0.000034	0.000363
SE	Epb41l3	chr17	+	69283884	69284007	0.000034	0.001959
5' ASS	Tmem214	chr5	+	30872686	30872904	0.000034	0.001332
RI	Tpp1	chr7	-	105750179	105751729	0.000034	0.000364
RI	Las1l	chrX	-	95953221	95955211	0.000034	0.000367
SE	Tmem57	chr4	-	134836766	134837034	0.000034	0.001980
SE	Ubfd1	chr7	+	122074466	122074557	0.000035	0.001986
RI	Uqcc1	chr2	-	155910356	155911836	0.000035	0.000370
RI	Ugt2b34	chr5	-	86889766	86892962	0.000036	0.000377
SE	Smarcc1	chr9	+	110236715	110236828	0.000036	0.002036
RI	Ppp1r12a	chr10	+	108261244	108262431	0.000036	0.000382
SE	Prpf4b	chr13	+	34893316	34893422	0.000037	0.002092
RI	Wdr1	chr5	-	38545673	38547091	0.000038	0.000395
RI	Stx5a	chr19	+	8743079	8743423	0.000038	0.000395
RI	Peli1	chr11	+	21140563	21142657	0.000039	0.000403
RI	Npr2	chr4	+	43643605	43644187	0.000039	0.000409
RI	Mettl26	chr17	+	25876660	25877169	0.000040	0.000416
3' ASS	Neo1	chr9	-	58906870	58907094	0.000040	0.002992
SE	Rbm25	chr12	+	83651645	83651750	0.000040	0.002253
SE	Spaca6	chr17	+	17837669	17837769	0.000041	0.002253
RI	Mms19	chr19	-	41945732	41947103	0.000041	0.000417
SE	Mkrn2os	chr6	-	115589344	115589394	0.000041	0.002284
RI	Mtx1	chr3	-	89210106	89210448	0.000041	0.000424
RI	Dusp1	chr17	-	26506984	26507719	0.000042	0.000425
RI	Fam214b	chr4	-	43035835	43037206	0.000042	0.000425
SE	Runx1t1	chr4	+	13771375	13771545	0.000043	0.002339
3' ASS	Chd6	chr2	-	161041984	161042132	0.000043	0.003095
RI	Nrbp1	chr5	+	31244165	31244559	0.000044	0.000443
RI	Acaa1a	chr9	+	119344465	119346568	0.000044	0.000445
RI	Cnbp	chr6	-	87845121	87845557	0.000045	0.000446
RI	Tbrg1	chr9	-	37654215	37655092	0.000045	0.000448
RI	Fibp	chr19	+	5463786	5464215	0.000045	0.000448
SE	Zfp952	chr17	+	32994554	32994703	0.000045	0.002439
SE	Cep164	chr9	-	45785777	45785921	0.000045	0.002439
SE	Nfu1	chr6	+	87016175	87016394	0.000046	0.002465
SE	Ppid	chr3	+	79596091	79596198	0.000046	0.002465
SE	Bptf	chr11	-	107095001	107095190	0.000047	0.002467

SE	Sema4d	chr13	-	51734506	51734570	0.000049	0.002541
RI	Tank	chr2	+	61643714	61644422	0.000049	0.000481
MXE	Epb41l2	chr10	+	25501577	25501631	0.000049	0.002588
3' ASS	Wbp1	chr6	-	83120184	83120876	0.000049	0.003442
SE	Pkm	chr9	+	59658998	59659153	0.000050	0.002573
SE	Sh2d5	chr4	+	138253611	138253689	0.000050	0.002573
RI	Mapk7	chr11	-	61492903	61494266	0.000050	0.000491
SE	Necap2	chr4	-	141071580	141071695	0.000050	0.002586
SE	Tnk2	chr16	+	32678668	32678713	0.000051	0.002586
RI	Zfas1	chr2	+	167065433	167065862	0.000051	0.000500
RI	A230050P20Rik	chr9	+	20869603	20871235	0.000051	0.000502
RI	Spag7	chr11	-	70664987	70665364	0.000052	0.000503
5' ASS	Ythdc1	chr5	+	86827954	86829479	0.000052	0.001989
RI	Sgsm3	chr15	+	81008288	81008908	0.000052	0.000504
RI	Trabd	chr15	+	89084902	89085446	0.000052	0.000504
RI	Rbm7	chr9	-	48488700	48490938	0.000052	0.000504
SE	Gripap1	chrX	+	7803593	7803686	0.000054	0.002719
SE	Mtmr6	chr14	+	60290705	60290819	0.000054	0.002724
SE	Eml5	chr12	-	98819860	98819869	0.000055	0.002724
5' ASS	Srrt	chr5	-	137301317	137302314	0.000055	0.002010
5' ASS	Traf2	chr2	-	25538844	25539084	0.000056	0.002010
RI	Gas5	chr1	+	161035751	161036014	0.000057	0.000541
SE	Hnrnpa2b1	chr6	-	51467390	51467426	0.000057	0.002828
SE	Golga4	chr9	+	118576222	118576356	0.000058	0.002864
RI	Sel1l	chr12	-	91813831	91814693	0.000059	0.000557
SE	Sgms1	chr19	-	32156569	32156647	0.000059	0.002876
RI	Ampd2	chr3	-	108080266	108080774	0.000059	0.000561
MXE	Ankrd6	chr4	-	32818604	32818709	0.000059	0.002588
RI	Gtf2h2	chr13	-	100486070	100486524	0.000060	0.000564
5' ASS	Thoc2	chrX	-	41813416	41814008	0.000060	0.002095
SE	Spaca6	chr17	+	17837664	17837771	0.000060	0.002901
MXE	Snap25	chr2	+	136769742	136769860	0.000060	0.002588
RI	Baz2b	chr2	-	59920294	59922217	0.000060	0.000567
SE	Mcf2l	chr8	+	13013914	13013998	0.000062	0.002988
RI	Snhg1	chr19	+	8724402	8724721	0.000062	0.000583
SE	Rnps1	chr17	+	24418037	24418218	0.000062	0.002988
RI	Pfkfb2	chr1	-	130697696	130698743	0.000063	0.000589
SE	Znhit1	chr5	-	136986125	136986255	0.000063	0.003018
RI	Nfx1	chr4	+	41015202	41016150	0.000064	0.000592
RI	Rpl3	chr15	-	80080895	80081783	0.000064	0.000595
SE	Arhgef1	chr7	+	24919183	24919351	0.000064	0.003045

SE	Tjap1	chr17	-	46264093	46264190	0.000065	0.003061
RI	Snx19	chr9	+	30432234	30433430	0.000065	0.000604
SE	Fnip2	chr3	-	79507838	79507928	0.000067	0.003150
SE	Golga2	chr2	+	32293314	32293395	0.000069	0.003170
SE	Dhx57	chr17	-	80278222	80278381	0.000069	0.003170
RI	Aup1	chr6	+	83055128	83055690	0.000069	0.000632
RI	Fam53a	chr5	-	33600355	33608227	0.000069	0.000632
5' ASS	Cyb561d1	chr3	-	108200626	108200834	0.000070	0.002391
SE	Clip1	chr5	-	123613067	123615335	0.000070	0.003218
RI	Odf2l	chr3	+	145126092	145127890	0.000071	0.000643
RI	Tmem214	chr5	+	30872686	30873144	0.000072	0.000655
RI	Aifm1	chrX	-	48483154	48484635	0.000074	0.000666
RI	Mrps2	chr2	+	28468242	28468950	0.000074	0.000666
RI	Iws1	chr18	+	32086272	32087152	0.000074	0.000666
SE	Zfp961	chr8	+	71966099	71966160	0.000074	0.003376
SE	Dock6	chr9	-	21829473	21829578	0.000075	0.003376
RI	Slc35f6	chr5	+	30655494	30656100	0.000075	0.000672
SE	Pramef8	chr4	+	143415348	143415413	0.000078	0.003517
RI	Bsdc1	chr4	+	129465229	129466990	0.000078	0.000697
3' ASS	Sorbs2	chr8	+	45823728	45823835	0.000079	0.005361
RI	Ubr2	chr17	-	46980319	46981442	0.000079	0.000706
SE	Spaca6	chr17	+	17837664	17837769	0.000080	0.003548
SE	Necap2	chr4	-	141071580	141071717	0.000080	0.003548
RI	Asb6	chr2	-	30824640	30825560	0.000080	0.000707
RI	Zcchc9	chr13	-	91800610	91801016	0.000080	0.000707
SE	Bsg	chr10	+	79708696	79709044	0.000081	0.003592
RI	Nob1	chr8	-	107416105	107417091	0.000081	0.000715
RI	Zer1	chr2	-	30108220	30110381	0.000083	0.000726
SE	Erbin	chr13	-	103830187	103830331	0.000083	0.003660
RI	Brd8	chr18	-	34608417	34609949	0.000083	0.000726
3' ASS	Dync2h1	chr9	-	7111424	7111592	0.000084	0.005467
RI	Parl	chr16	-	20285764	20287069	0.000084	0.000730
RI	Ankrd54	chr15	-	79054495	79055416	0.000085	0.000736
SE	Ddx31	chr2	+	28846762	28846871	0.000085	0.003695
SE	Kctd12b	chrX	-	153695623	153695678	0.000085	0.003695
5' ASS	Arhgap26	chr18	+	39357552	39357828	0.000086	0.002840
RI	Tpp2	chr1	+	44001318	44003000	0.000087	0.000748
RI	Hcfc1	chrX	-	73946709	73947050	0.000089	0.000761
5' ASS	Hdac10	chr15	-	89126629	89126706	0.000089	0.002840
MXE	Fam184a	chr10	-	53634301	53634395	0.000090	0.003526
RI	Atp11b	chr3	+	35837485	35839214	0.000090	0.000770

SE	Arhgef1	chr7	+	24918519	24919351	0.000091	0.003920
5' ASS	Zmynd8	chr2	-	165807518	165807929	0.000091	0.002840
5' ASS	Ankrd16	chr2	+	11783640	11783756	0.000092	0.002840
RI	Srp19	chr18	+	34331749	34334364	0.000092	0.000782
RI	Tbrg1	chr9	-	37652614	37653169	0.000092	0.000782
RI	Tfe3	chrX	+	7770994	7772596	0.000092	0.000782
SE	Nubp1	chr16	+	10421870	10421985	0.000093	0.003990
RI	Slc25a22	chr7	-	141432020	141432753	0.000096	0.000811
SE	Ttc13	chr8	-	124699655	124699721	0.000097	0.004152
RI	Ppid	chr3	+	79595201	79597873	0.000098	0.000825
SE	Anapc16	chr10	-	60002417	60002467	0.000099	0.004212
RI	Wdr77	chr3	+	105965764	105966425	0.000100	0.000837
RI	Hps5	chr7	-	46774813	46775951	0.000102	0.000851
MXE	Gria3	chrX	+	41654809	41654924	0.000102	0.003681
3' ASS	Ddx49	chr8	-	70294686	70294835	0.000104	0.006631
RI	Fxyd5	chr7	-	31032729	31033018	0.000107	0.000888
RI	Mfsd11	chr11	+	116863905	116865956	0.000107	0.000890
SE	Atp11b	chr3	+	35843570	35843696	0.000109	0.004603
5' ASS	Ankrd16	chr2	+	11779676	11779897	0.000110	0.003309
SE	Taz	chrX	+	74288983	74289019	0.000110	0.004611
SE	Sidt2	chr9	-	45946505	45946568	0.000111	0.004628
RI	Fam102a	chr2	+	32558317	32560155	0.000112	0.000927
SE	Tmem192	chr8	+	64964197	64964532	0.000113	0.004680
RI	Hdac3	chr18	-	37941378	37941766	0.000113	0.000931
RI	Hsd3b7	chr7	+	127802235	127803802	0.000114	0.000935
RI	1110032A03Rik	chr9	-	50764223	50764926	0.000115	0.000940
RI	Nudt13	chr14	+	20306231	20307799	0.000116	0.000940
RI	Psmd5	chr2	-	34856447	34857829	0.000116	0.000940
MXE	Pkm	chr9	+	59675028	59675195	0.000116	0.003863
RI	Nbr1	chr11	+	101574998	101575416	0.000117	0.000947
SE	Ap3m1	chr14	-	21044652	21044824	0.000118	0.004851
RI	Psmd4	chr3	-	95032704	95033546	0.000118	0.000949
SE	Srrm1	chr4	-	135345195	135345477	0.000120	0.004907
SE	Ncapd3	chr9	+	27030622	27030830	0.000120	0.004907
RI	Hmgn3	chr9	-	83109947	83111106	0.000120	0.000966
5' ASS	Fam229b	chr10	-	39133822	39133900	0.000122	0.003544
SE	Pax6	chr2	+	105684886	105685429	0.000123	0.003544
RI	Traf4	chr11	-	78161286	78161622	0.000124	0.000989
SE	Lpgat1	chr1	+	191718428	191718597	0.000124	0.005024
RI	Zswim8	chr14	+	20722459	20723006	0.000125	0.000996
RI	Metap2	chr10	-	93879539	93880138	0.000126	0.001002

3' ASS	Nfya	chr17	-	48392339	48392506	0.000127	0.007833
RI	Dpp7	chr2	-	25354892	25355721	0.000129	0.001022
RI	Gas5	chr1	+	161037504	161038037	0.000130	0.001028
RI	Prpf18	chr2	-	4637147	4638995	0.000132	0.001040
3' ASS	Trip10	chr17	+	57262353	57262472	0.000132	0.007833
RI	Med25	chr7	-	44879390	44880104	0.000134	0.001047
3' ASS	Helq	chr5	-	100768281	100768699	0.000134	0.007833
RI	Dcp1b	chr6	+	119214769	119217979	0.000134	0.001047
SE	Dtx2	chr5	+	136026469	136026607	0.000135	0.005464
SE	2010111I01Rik	chr13	+	63300604	63300793	0.000137	0.005490
RI	Zfp384	chr6	+	125024032	125025082	0.000138	0.001074
RI	E130309D02Rik	chr5	-	143311710	143314399	0.000138	0.001074
SE	Bptf	chr11	-	107086711	107086897	0.000139	0.005545
3' ASS	Jkamp	chr12	+	72089314	72089406	0.000139	0.007888
3' ASS	Zzz3	chr3	+	152449012	152449128	0.000142	0.007888
5' ASS	Katnal2	chr18	-	77011983	77012104	0.000142	0.004004
MXE	Pfkp	chr13	-	6593100	6593253	0.000142	0.004390
RI	Abtb1	chr6	-	88840709	88840900	0.000144	0.001115
SE	Aamdc	chr7	-	97575587	97575691	0.000146	0.005779
SE	Rdh13	chr7	-	4443736	4443771	0.000146	0.005779
5' ASS	Selenoi	chr5	+	30257711	30257820	0.000147	0.004075
SE	Nt5c2	chr19	-	46898327	46898521	0.000148	0.005810
SE	Nt5c3	chr6	-	56897774	56897829	0.000150	0.005847
3' ASS	Ehbp1	chr11	-	22146462	22146700	0.000151	0.008219
5' ASS	Fam229b	chr10	-	39133811	39133900	0.000153	0.004142
RI	Brd9	chr13	+	73941943	73942806	0.000157	0.001208
3' ASS	Dnajb2	chr1	+	75239107	75239333	0.000158	0.008384
RI	Ppia	chr11	+	6418100	6418314	0.000158	0.001217
RI	Adam17	chr12	-	21329053	21330163	0.000159	0.001217

Table Appendix-2. List of genes significantly alternatively spliced between human alpha and beta cells.

Splicing Event	Gene	Chr	Strand	Exon_Start	Exon_End	PValue	FDR
5' ASS	POPDC3	chr6	-	105161424	105162160	0.000000	0.000000
RI	LLGL1	chr17	+	18235469	18236760	0.000000	0.000000
MXE	SLC7A2	chr8	+	17554559	17554699	0.000000	0.000030
5' ASS	CHD5	chr1	-	6130147	6130328	0.000000	0.000086
SE	CAST	chr5	+	96740744	96740783	0.000000	0.001277
SE	ZNF35	chr3	+	44652556	44652701	0.000000	0.001633
SE	CCNT1	chr12	-	48695998	48696162	0.000000	0.001633
RI	NDRG2	chr14	-	21017986	21018814	0.000000	0.000925
SE	PEG3	chr19	-	56833289	56833420	0.000001	0.003701
SE	SPATS2L	chr2	+	200308886	200309090	0.000001	0.003701
3'ASS	AHDC1	chr1	-	27553073	27553304	0.000001	0.005376
SE	RMDN2	chr2	+	37951232	37952201	0.000001	0.004591
SE	LMF1	chr16	-	855106	856055	0.000002	0.005153
3'ASS	AHDC1	chr1	-	27553073	27553290	0.000003	0.006867
MXE	SLC3A2	chr11	+	62882504	62882618	0.000005	0.004869
RI	LDHD	chr16	-	75114826	75115339	0.000006	0.006732
RI	NT5C3B	chr17	-	41835587	41835957	0.000007	0.006732
SE	DOCK8	chr9	+	365590	366552	0.000008	0.019597
SE	TRANK1	chr3	-	36850153	36851417	0.000009	0.019597
RI	AKR1C2	chr10	-	4999199	5000666	0.000009	0.007013
3'ASS	SYNGAP1	chr6	+	33440552	33440965	0.000010	0.014692
5' ASS	PYROXD2	chr10	-	98410612	98410958	0.000011	0.008634
5' ASS	ZNF598	chr16	-	2000390	2000481	0.000012	0.008634
SE	ARFGAP2	chr11	-	47172709	47172751	0.000014	0.026272
SE	PUF60	chr8	-	143820665	143820716	0.000014	0.026272
SE	TRIM36	chr5	-	115163517	115163752	0.000021	0.034494
RI	SYTL1	chr1	+	27349971	27350485	0.000025	0.016503
SE	INO80E	chr16	+	30001211	30004657	0.000026	0.039020
SE	AHDC1	chr1	-	27558304	27558530	0.000027	0.039020
SE	ZNHIT1	chr7	+	101218318	101218485	0.000032	0.043454
SE	AC055758.2	chr3	+	145946403	145946621	0.000039	0.049608
RI	NUDT8	chr11	-	67628918	67629921	0.000042	0.021870
RI	SSR2	chr1	-	156011809	156012565	0.000045	0.021870
SE	FAAP20	chr1	-	2192646	2192975	0.000047	0.055057
RI	RABGGTB	chr1	+	75789951	75790274	0.000064	0.027519
5' ASS	MPRIP	chr17	+	17137915	17138429	0.000067	0.038387
MXE	PEG3	chr19	-	56829782	56829921	0.000079	0.047095
SE	ZNHIT1	chr7	+	101218318	101221796	0.000087	0.093334

SE	LINC00963	chr9	+	129493477	129493595	0.000092	0.093334
SE	DDX55	chr12	+	123607431	123607523	0.000093	0.093334
5' ASS	MFF	chr2	+	227328677	227328899	0.000100	0.047621
RI	HMG20B	chr19	+	3573691	3574586	0.000102	0.039554
SE	DNAJB5	chr9	+	34990347	34990611	0.000135	0.129500
5' ASS	ALS2CR12	chr2	-	201348099	201348302	0.000137	0.055487
RI	MAP4K2	chr11	-	64791908	64792422	0.000137	0.045662
3'ASS	SRSF5	chr14	+	69769510	69770540	0.000143	0.152612
RI	DHRS1	chr14	-	24291138	24291625	0.000156	0.045662
RI	YDJC	chr22	-	21629307	21629761	0.000167	0.045662
RI	QTRT1	chr19	+	10712160	10712628	0.000171	0.045662
RI	IL17RC	chr3	+	9923880	9924291	0.000176	0.045662
SE	SYNRG	chr17	-	37579131	37579434	0.000177	0.161429
SE	AMDHD2	chr16	+	2527572	2527615	0.000189	0.165546
5' ASS	PYROXD2	chr10	-	98390870	98391082	0.000200	0.071186
RI	KRI1	chr19	-	10557551	10557895	0.000225	0.049016
SE	AC132008.2	chr3	-	197624471	197624720	0.000229	0.192113
RI	PIGO	chr9	-	35091239	35092767	0.000238	0.049016
SE	NDUFS8	chr11	+	68031532	68031710	0.000247	0.198424
RI	SPAG5	chr17	-	28578217	28578528	0.000250	0.049016
RI	HEXDC	chr17	+	82436666	82437363	0.000258	0.049016
RI	SLC25A39	chr17	-	44321057	44321558	0.000264	0.049016
RI	HAX1	chr1	+	154273335	154273961	0.000264	0.049016
SE	FAAP20	chr1	-	2192844	2192975	0.000318	0.242233
SE	PAPSS2	chr10	+	87721755	87721770	0.000325	0.242233
3'ASS	RELA	chr11	-	65662178	65662347	0.000354	0.260617
RI	CARD19	chr9	+	93110567	93112289	0.000361	0.062167
3'ASS	DACH2	chrX	+	86827765	86832602	0.000366	0.260617
RI	TRIM69	chr15	+	44756367	44758854	0.000367	0.062167
SE	GNAS	chr20	+	58891581	58891865	0.000380	0.268411
SE	CIDEB	chr14	-	24310406	24310548	0.000387	0.268411
RI	PELP1	chr17	-	4675273	4675884	0.000453	0.073523
3'ASS	CDK18	chr1	+	205530185	205530349	0.000456	0.278455
MXE	NSMAF	chr8	-	58637233	58637389	0.000464	0.208562
5' ASS	LCLAT1	chr2	+	30447235	30447404	0.000511	0.146769
5' ASS	DDX47	chr12	+	12823202	12824016	0.000516	0.146769
RI	LIN37	chr19	+	35753086	35754157	0.000520	0.078843
SE	SETD5	chr3	+	9433369	9433562	0.000534	0.339625
SE	NGLY1	chr3	-	25732318	25732483	0.000539	0.339625
RI	YBX3	chr12	-	10701959	10704148	0.000549	0.078843
SE	SDR42E1	chr16	-	82007640	82007868	0.000551	0.339625

3'ASS	CLUH	chr17	-	2696433	2696538	0.000556	0.297313
RI	SKIV2L	chr6	+	31969279	31969751	0.000560	0.078843
RI	CRACR2B	chr11	+	830865	831295	0.000567	0.078843
SE	WSB1	chr17	+	27301817	27301956	0.000569	0.339625
SE	AC006001.3	chr7	-	66576754	66576932	0.000574	0.339625
SE	EPS8L2	chr11	+	706595	706801	0.000625	0.359231
5' ASS	IRF3	chr19	-	49661449	49662328	0.000697	0.179348
RI	PGAP2	chr11	+	3823882	3824376	0.000706	0.094770
SE	COL9A2	chr1	-	40315237	40315465	0.000714	0.389801
SE	NAA16	chr13	+	41362730	41362873	0.000717	0.389801
5' ASS	HDAC9	chr7	+	18495745	18496023	0.000756	0.179348
RI	PNKP	chr19	-	49864178	49864403	0.000781	0.099061
SE	SRSF2	chr17	-	76735771	76735875	0.000793	0.419324
RI	ZSWIM7	chr17	-	15977578	15977913	0.000812	0.099061
RI	ECEL1	chr2	-	232484980	232485267	0.000847	0.099061
RI	CPSF7	chr11	-	61420469	61421608	0.000849	0.099061
SE	DOCK8	chr9	+	369221	369447	0.000865	0.445805
RI	RENBP	chrX	-	153944079	153944418	0.000878	0.099061
RI	ABCA2	chr9	-	137011406	137011749	0.000890	0.099061
SE	ABHD18	chr4	+	127989720	127989821	0.000928	0.466416
RI	ITGA7	chr12	-	55694777	55695637	0.000969	0.100060
RI	APBA3	chr19	-	3751188	3751553	0.000982	0.100060
RI	DMKN	chr19	-	35500377	35500580	0.000985	0.100060
RI	SAT2	chr17	-	7627362	7627802	0.001002	0.100060
SE	PIGQ	chr16	+	581198	581341	0.001015	0.481191
SE	SUGP2	chr19	-	18993647	18993740	0.001017	0.481191
SE	FRS2	chr12	+	69480298	69480359	0.001029	0.481191
MXE	MROH7	chr1	+	54679890	54680045	0.001045	0.375727
5' ASS	PTPRZ1	chr7	+	122010333	122013889	0.001059	0.231807
RI	SRSF5	chr14	+	69769181	69770540	0.001091	0.102629
RI	MAP4K2	chr11	-	64792171	64792422	0.001095	0.102629
RI	CDK10	chr16	+	89693273	89694232	0.001107	0.102629
SE	SUGP2	chr19	-	18993637	18993740	0.001167	0.533168
RI	MVD	chr16	-	88657435	88658029	0.001321	0.119284
RI	PNKP	chr19	-	49861607	49861881	0.001348	0.119284
SE	POLDIP3	chr22	-	42601969	42602107	0.001376	0.614610
MXE	HAUS2	chr15	+	42558197	42558290	0.001399	0.418994
5' ASS	ABHD11	chr7	-	73737220	73737735	0.001441	0.285469
5' ASS	CENPX	chr17	-	82019509	82019694	0.001505	0.285469
RI	SIRT6	chr19	-	4175027	4175937	0.001518	0.130615
RI	UNC119	chr17	-	28546707	28547849	0.001543	0.130615

RI	MAU2	chr19	+	19344848	19345369	0.001704	0.141213
RI	DRG2	chr17	+	18099016	18100435	0.001801	0.146137
SE	ITGA9-AS1	chr3	-	37818679	37818930	0.001814	0.772581
SE	LIPC	chr15	+	58430826	58431151	0.001831	0.772581
SE	ZNF595	chr4	+	60057	60153	0.001850	0.772581
RI	CARD8	chr19	-	48232452	48234543	0.001863	0.148072
SE	WSB1	chr17	+	27301880	27301956	0.001883	0.772581
RI	CSK	chr15	+	74800680	74800922	0.002003	0.154155
SE	FBXO44	chr1	+	11658532	11658628	0.002031	0.813243
SE	EPB41L1	chr20	+	36195328	36195364	0.002063	0.813243
RI	ERBB3	chr12	+	56094098	56094556	0.002068	0.154155
RI	CYP4F3	chr19	+	15647051	15647324	0.002086	0.154155
RI	CDK18	chr1	+	205529520	205530349	0.002098	0.154155
SE	FAM86EP	chr4	-	3947143	3948220	0.002210	0.849500
SE	LIAS	chr4	+	39467517	39468427	0.002240	0.849500
MXE	SLC25A45	chr11	-	65379375	65379561	0.002303	0.591269
RI	THOP1	chr19	+	2807441	2808444	0.002335	0.168367
RI	RNH1	chr11	-	499014	499999	0.002405	0.170240
SE	NME4	chr16	+	398207	398394	0.002554	0.950720
RI	MGAT4B	chr5	-	179801333	179801694	0.002555	0.176169
SE	RNF135	chr17	+	30987943	30988106	0.002608	0.953254
RI	SBF1	chr22	-	50467531	50467923	0.002622	0.176169
RI	CXXC1	chr18	-	50285748	50286257	0.002668	0.176169
RI	TPP1	chr11	-	6619195	6619413	0.002669	0.176169
5' ASS	KL	chr13	+	33055046	33055365	0.002678	0.449524
5' ASS	HMGA1	chr6	+	34240736	34240915	0.002686	0.449524
SE	LINC00963	chr9	+	129493418	129493595	0.002698	0.962583
SE	ATG9A	chr2	-	219229009	219229174	0.002730	0.962583
RI	NECAB3	chr20	-	33659884	33660395	0.002753	0.178652
RI	ARAP1	chr11	-	72712437	72713243	0.002845	0.181631
SE	APOPT1	chr14	+	103571622	103571820	0.002866	0.962583
SE	LINC00963	chr9	+	129493415	129493595	0.002895	0.962583
3'ASS	MGRN1	chr16	+	4686258	4690974	0.002906	1.000000
SE	ATP6V0A1	chr17	+	42510058	42510183	0.002920	0.962583
SE	RHOC	chr1	-	112705099	112706252	0.002921	0.962583
SE	OSBPL9	chr1	+	51756319	51756358	0.003101	1.000000
RI	DOCK8	chr9	+	368017	370300	0.003153	0.198038
5' ASS	NAA16	chr13	+	41362030	41362873	0.003251	0.497012
RI	MZF1	chr19	-	58570343	58571429	0.003255	0.201189
SE	MEG8	chr14	+	100977604	100977722	0.003267	1.000000
SE	MRPL9	chr1	-	151761450	151761552	0.003273	1.000000

5' ASS	VRK3	chr19	-	50025266	50025372	0.003319	0.497012
3'ASS	CACNA1C	chr12	+	2651573	2651768	0.003354	1.000000
RI	SCARB1	chr12	-	124782682	124786503	0.003371	0.205081
3'ASS	CLK4	chr5	-	178619837	178620673	0.003424	1.000000
SE	PBRM1	chr3	-	52682091	52682233	0.003501	1.000000
SE	ZBTB37	chr1	+	173873466	173873566	0.003521	1.000000
SE	SENP7	chr3	-	101501069	101501119	0.003551	1.000000
RI	GPAA1	chr8	+	144083388	144083861	0.003616	0.216605
5' ASS	C7orf43	chr7	-	100157292	100157459	0.003636	0.517286
3'ASS	ZFP30	chr19	-	37643264	37643363	0.003644	1.000000
SE	PEX5	chr12	+	7190021	7190216	0.003759	1.000000
MXE	BORA	chr13	+	72728925	72729093	0.003816	0.857257
SE	SYNRG	chr17	-	37520178	37520214	0.003819	1.000000
3'ASS	ARL4A	chr7	+	12688165	12689176	0.003831	1.000000
5' ASS	THAP5	chr7	-	108565829	108566022	0.003885	0.526323
SE	POLDIP3	chr22	-	42601969	42602056	0.003895	1.000000
RI	CNBP	chr3	-	129171078	129171538	0.003950	0.233047
SE	CHD2	chr15	+	92901166	92901299	0.003999	1.000000
SE	TOPORS	chr9	-	32550773	32550968	0.004095	1.000000
SE	MUC20	chr3	+	195722500	195722818	0.004119	1.000000
5' ASS	ALMS1	chr2	+	73489633	73491498	0.004188	0.541642
RI	SHISA5	chr3	-	48467875	48469573	0.004199	0.241789
RI	IDH3G	chrX	-	153785765	153786272	0.004222	0.241789
MXE	ISYNA1	chr19	-	18436683	18436877	0.004340	0.866467
SE	CEP164	chr11	+	117382795	117382942	0.004442	1.000000
SE	SMCHD1	chr18	+	2722518	2722663	0.004524	1.000000
SE	SRSF11	chr1	+	70231858	70231975	0.004551	1.000000
SE	PKM	chr15	-	72221167	72221284	0.004644	1.000000
3'ASS	UBE2D3	chr4	-	102826484	102826844	0.004690	1.000000
3'ASS	SLC2A11	chr22	+	23877384	23877869	0.004740	1.000000
SE	MEIS2	chr15	-	36894716	36894812	0.004764	1.000000
RI	ZFPL1	chr11	+	65085114	65086608	0.004965	0.266038
RI	SPNS2	chr17	+	4536075	4536426	0.004991	0.266038
RI	SPAG4	chr20	+	35618449	35618720	0.005024	0.266038
5' ASS	PSTK	chr10	+	122983271	122983514	0.005069	0.602466
5' ASS	SH3BP1	chr22	+	37642538	37642737	0.005082	0.602466
RI	SARM1	chr17	+	28388173	28388539	0.005168	0.266038
RI	RCHY1	chr4	-	75490580	75491782	0.005179	0.266038
RI	NEK9	chr14	-	75105949	75106702	0.005191	0.266038
RI	FDXR	chr17	-	74863895	74864347	0.005205	0.266038
RI	TMEM176A	chr7	+	150801535	150802325	0.005211	0.266038

RI	WDR81	chr17	+	1735571	1736218	0.005261	0.266038
5' ASS	MAPK8IP3	chr16	+	1761223	1761305	0.005336	0.607272
SE	ZNF772	chr19	-	57475659	57475786	0.005431	1.000000
MXE	AC022784.1	chr8	+	9168110	9169177	0.005498	0.987942
5' ASS	ITGA7	chr12	-	55700565	55701154	0.005683	0.621820
SE	NDRG4	chr16	+	58500987	58501077	0.005782	1.000000
RI	PQBP1	chrX	+	48902232	48902795	0.005892	0.292040
SE	FAM126A	chr7	-	22946951	22947247	0.005897	1.000000
RI	RBM6	chr3	+	50061461	50062108	0.005994	0.292040
RI	LARGE2	chr11	+	45926038	45926347	0.006220	0.292040
RI	REX1BD	chr19	+	18588994	18589683	0.006274	0.292040
RI	BTN3A2	chr6	+	26373275	26373413	0.006292	0.292040
RI	NBEAL2	chr3	+	47007240	47007697	0.006333	0.292040
RI	GTF2H4	chr6	+	30911683	30912146	0.006344	0.292040
RI	GPAT2	chr2	-	96024772	96025603	0.006375	0.292040
SE	EHBP1	chr2	+	62949093	62949162	0.006439	1.000000
SE	FAM208B	chr10	+	5709529	5709663	0.006455	1.000000
SE	SYNGR3	chr16	+	1990498	1990672	0.006477	1.000000
SE	TDRD9	chr14	+	104049864	104049937	0.006847	1.000000
SE	ETFA	chr15	-	76295590	76295737	0.006861	1.000000
5' ASS	LZTR1	chr22	+	20990385	20990969	0.006867	0.723582
SE	MXD4	chr4	-	2255154	2255351	0.007124	1.000000
SE	REPIN1	chr7	+	150370713	150370884	0.007254	1.000000
MXE	SGSH	chr17	-	80214614	80214765	0.007397	1.000000
SE	DCP1A	chr3	-	53304176	53304290	0.007496	1.000000
MXE	CRTC3	chr15	+	90614452	90614488	0.007633	1.000000
SE	RFC5	chr12	+	118017689	118017815	0.008004	1.000000
SE	AC156455.1	chr12	+	122064138	122064433	0.008004	1.000000
SE	PPIP5K2	chr5	+	103135992	103136114	0.008051	1.000000
3'ASS	WDR63	chr1	+	85093773	85094555	0.008156	1.000000
RI	HNF1A	chr12	+	120999267	120999627	0.008207	0.363403
RI	C1orf35	chr1	-	228102309	228102560	0.008215	0.363403
RI	RNF123	chr3	+	49712478	49713587	0.008278	0.363403
RI	GRM4	chr6	-	34058973	34062028	0.008306	0.363403
RI	DMKN	chr19	-	35510183	35510510	0.008519	0.364278
RI	HPN	chr19	+	35041721	35042522	0.008609	0.364278
RI	LRRC75A-AS1	chr17	+	16439326	16439703	0.008644	0.364278
RI	UNC93B1	chr11	-	68003021	68003798	0.008700	0.364278
RI	INO80E	chr16	+	30001211	30004657	0.008844	0.366348
5' ASS	ELMOD3	chr2	+	85371085	85371562	0.008911	0.905428
SE	FAM131A	chr3	+	184338386	184338529	0.008912	1.000000

RI	CTSA	chr20	+	45891568	45892027	0.009008	0.368363
SE	CAPNS1	chr19	+	36140357	36140511	0.009054	1.000000
RI	YDJC	chr22	-	21629307	21629761	0.009081	0.368363
SE	RBM5	chr3	+	50110378	50110463	0.009128	1.000000
SE	DZIP1L	chr3	-	138071642	138071835	0.009174	1.000000
RI	SH3BP2	chr4	+	2831922	2832412	0.009237	0.370797
MXE	KLHDC8A	chr1	-	205344133	205344458	0.009267	1.000000
3'ASS	MVD	chr16	-	88656104	88656561	0.009291	1.000000
SE	MPHOSPH9	chr12	-	123203249	123203375	0.009407	1.000000
5' ASS	MAP3K3	chr17	+	63685516	63685590	0.009426	0.924722
RI	TMEM175	chr4	+	947708	948154	0.009463	0.371655
RI	TRPM4	chr19	+	49188640	49189091	0.009477	0.371655
SE	MZF1	chr19	-	58572549	58572719	0.009490	1.000000
SE	FAM86EP	chr4	-	3948089	3948220	0.009572	1.000000
SE	ABHD11	chr7	-	73737220	73737360	0.009589	1.000000
RI	PTK7	chr6	+	43142171	43143620	0.009637	0.371655
RI	INTS3	chr1	+	153772339	153772711	0.009640	0.371655
SE	BAZ2A	chr12	-	56630775	56630865	0.009663	1.000000
3'ASS	CCDC74A	chr2	+	131530578	131530827	0.009736	1.000000
RI	CFAP70	chr10	-	73293262	73297173	0.009895	0.375043
RI	POMT1	chr9	+	131509902	131510415	0.009920	0.375043
5' ASS	FUZ	chr19	-	49812446	49812736	0.009989	0.941493
3'ASS	INTS11	chr1	-	1313731	1314105	0.010171	1.000000
SE	CTDP1	chr18	+	79728906	79729069	0.010172	1.000000
SE	ZNF708	chr19	-	21327959	21328141	0.010189	1.000000
3'ASS	ZNF711	chrX	+	85246929	85247651	0.010209	1.000000
SE	GNE	chr9	-	36236831	36236984	0.010220	1.000000
5' ASS	BUD23	chr7	+	73683766	73684000	0.010259	0.941493
SE	SUOX	chr12	+	55997614	55997723	0.010273	1.000000
RI	HDAC10	chr22	-	50246297	50246735	0.010476	0.392245
RI	KANK1	chr9	+	732377	734835	0.010684	0.393799
RI	REX1BD	chr19	+	18588994	18589683	0.010720	0.393799
3'ASS	STRADA	chr17	-	63704483	63704908	0.010723	1.000000
SE	MYH14	chr19	+	50224153	50224177	0.010834	1.000000
RI	NELFCD	chr20	+	58993623	58994239	0.010848	0.394783
5' ASS	NT5C2	chr10	-	103100662	103100943	0.011011	0.961719
RI	MMS19	chr10	-	97461495	97461896	0.011076	0.396757
MXE	MFSD8	chr4	-	127939852	127939997	0.011084	1.000000
RI	RAB26	chr16	+	2152819	2153222	0.011106	0.396757
3'ASS	TGFBR3	chr1	-	91719892	91720230	0.011145	1.000000
SE	UBE2D4	chr7	+	43942953	43943031	0.011220	1.000000

5' ASS	WDR91	chr7	-	135197043	135198151	0.011280	0.961719
SE	C19orf48	chr19	-	50802453	50802614	0.011404	1.000000
RI	ELMOD3	chr2	+	85355075	85357252	0.011477	0.404187
5' ASS	CCDC7	chr10	+	32446139	32446631	0.011493	0.961719
MXE	DNM2	chr19	+	10796060	10796199	0.011501	1.000000
SE	INO80E	chr16	+	30001211	30001530	0.011502	1.000000
SE	AMPD2	chr1	+	109620913	109621266	0.011524	1.000000
SE	ZNF415	chr19	-	53123500	53123655	0.011574	1.000000
RI	C1orf159	chr1	-	1084352	1084506	0.011636	0.404187
RI	TPCN1	chr12	+	113288162	113288847	0.011784	0.404187
RI	TRABD	chr22	+	50197822	50198186	0.011815	0.404187
SE	FAM126A	chr7	-	22947032	22947247	0.011826	1.000000
RI	NKTR	chr3	+	42619663	42621516	0.011894	0.404187
RI	RUSC1	chr1	+	155325863	155327132	0.011937	0.404187
SE	PEG3	chr19	-	56833289	56833420	0.012089	1.000000
RI	TANC1	chr2	+	159219237	159219867	0.012440	0.414513
SE	BPHL	chr6	+	3119284	3119512	0.012557	1.000000
RI	MRPL4	chr19	+	10252396	10252701	0.012561	0.414513
RI	RNF213	chr17	+	80381546	80383070	0.012683	0.414513
RI	ANKRD13D	chr11	+	67291475	67291746	0.012690	0.414513
MXE	ZNF415	chr19	-	53115209	53115307	0.012746	1.000000
SE	MTMR14	chr3	+	9697710	9697866	0.012757	1.000000
RI	ARSA	chr22	-	50627555	50628170	0.012824	0.414513
RI	ZNF662	chr3	+	42908021	42908909	0.012880	0.414513
RI	DDX51	chr12	-	132141497	132141956	0.013109	0.418039
SE	APOPT1	chr14	+	103563214	103563379	0.013152	1.000000
RI	NCKIPSD	chr3	-	48678876	48679183	0.013240	0.418039
SE	RRNAD1	chr1	+	156734008	156734493	0.013249	1.000000
RI	REG3A	chr2	-	79159329	79159753	0.013312	0.418039
RI	DXO	chr6	-	31970605	31971147	0.013448	0.418944
RI	KLHDC4	chr16	-	87755193	87756477	0.013725	0.423430
RI	AL096870.2	chr14	-	24215551	24215797	0.013876	0.423430
RI	KLHL35	chr11	-	75425392	75426638	0.013919	0.423430
3'ASS	PLD3	chr19	+	40365552	40365930	0.013933	1.000000
3'ASS	INO80E	chr16	+	30001211	30001530	0.013943	1.000000
5' ASS	PLEKHH3	chr17	-	42670993	42671130	0.013948	1.000000
3'ASS	KRT40	chr17	-	40980989	40981324	0.014072	1.000000
3'ASS	ACSL4	chrX	-	109683135	109683666	0.014079	1.000000
5' ASS	MVD	chr16	-	88657734	88658029	0.014115	1.000000
5' ASS	RASGRF1	chr15	-	79006176	79006434	0.014324	1.000000
RI	SFI1	chr22	+	31618114	31618547	0.014333	0.429386

RI	NOL3	chr16	+	67174161	67174944	0.014335	0.429386
SE	NAB1	chr2	+	190655976	190656153	0.014410	1.000000
SE	CPNE1	chr20	-	35658929	35658982	0.014450	1.000000
SE	GGCX	chr2	-	85560814	85560985	0.014555	1.000000
RI	WDR73	chr15	-	84645470	84646348	0.014631	0.434903
SE	SEC31A	chr4	-	82830936	82830975	0.014653	1.000000
3'ASS	APBB1	chr11	-	6401573	6401982	0.014917	1.000000
RI	KDM2B	chr12	-	121439856	121440977	0.014925	0.440289
SE	ARL13B	chr3	+	93996547	93996684	0.014937	1.000000
SE	ZNRD1ASP	chr6	-	30057090	30057717	0.014990	1.000000
SE	GULP1	chr2	+	188582340	188582565	0.015223	1.000000
RI	BEST1	chr11	+	61958145	61959578	0.015229	0.445868
RI	DDIT3	chr12	-	57517268	57517753	0.015654	0.454898
MXE	CTDP1	chr18	+	79728906	79729069	0.015655	1.000000
SE	EPHA10	chr1	-	37722005	37722296	0.015685	1.000000
SE	TEAD1	chr11	+	12675408	12675561	0.015719	1.000000
SE	BTBD3	chr20	+	11919085	11919176	0.015783	1.000000
SE	CTPS1	chr1	+	40980885	40981043	0.015872	1.000000
5' ASS	EPOR	chr19	-	11382846	11383232	0.015875	1.000000
SE	APCDD1L-DT	chr20	+	58519503	58519625	0.015886	1.000000
SE	MPHOSPH9	chr12	-	123227462	123227572	0.016277	1.000000
5' ASS	MTO1	chr6	+	73479935	73480367	0.016539	1.000000
SE	FAHD2A	chr2	+	95405276	95405402	0.016542	1.000000
RI	CC2D1B	chr1	-	52357780	52358461	0.016822	0.485213
RI	AAAS	chr12	-	53314297	53314849	0.017051	0.485961
RI	AC104452.1	chr3	-	49419709	49420342	0.017166	0.485961
RI	RPL10A	chr6	+	35468946	35469529	0.017222	0.485961
SE	ZNF860	chr3	+	31986247	31986409	0.017277	1.000000
RI	PMM1	chr22	-	41583958	41584372	0.017443	0.488658
SE	NR4A1	chr12	+	52052473	52052645	0.017493	1.000000
MXE	BTN2A2	chr6	+	26385014	26385362	0.017716	1.000000
RI	STAT2	chr12	-	56354777	56355351	0.017794	0.494625
SE	EBPL	chr13	-	49669776	49669846	0.017803	1.000000
RI	SGSM3	chr22	+	40408932	40409372	0.017973	0.494625
RI	CCDC153	chr11	-	119193150	119193885	0.018037	0.494625
SE	SUOX	chr12	+	55999332	55999458	0.018041	1.000000
SE	SGCE	chr7	-	94644605	94644667	0.018170	1.000000
RI	CFB	chr6	+	31947944	31948512	0.018287	0.494938
RI	ASL	chr7	+	66089090	66089335	0.018377	0.494938
3'ASS	C1orf159	chr1	-	1090352	1091374	0.018429	1.000000
RI	CFAP410	chr21	-	44330462	44332014	0.018430	0.494938

RI	NUBP2	chr16	+	1786536	1787831	0.018614	0.496470
SE	KIF21A	chr12	-	39318072	39318201	0.018671	1.000000
SE	LUCAT1	chr5	-	91311645	91312413	0.018810	1.000000
RI	FNDC4	chr2	-	27493929	27494466	0.018812	0.498327
SE	FAM228B	chr2	+	24167626	24167683	0.018829	1.000000
3'ASS	DAZAP1	chr19	+	1434736	1435687	0.019041	1.000000
SE	TRIM33	chr1	-	114397939	114397990	0.019171	1.000000
RI	AC104452.1	chr3	-	49419259	49420342	0.019215	0.505567
3'ASS	ARL13B	chr3	+	93996547	93996684	0.019396	1.000000
RI	TPRN	chr9	-	137191616	137192365	0.019445	0.507663
RI	KIF22	chr16	+	29799263	29799781	0.019556	0.507663
3'ASS	BTN2A1	chr6	+	26467718	26469637	0.019752	1.000000
RI	PPP4R1L	chr20	-	58245782	58247506	0.020185	0.520527
SE	CLIP2	chr7	+	74372931	74373036	0.020194	1.000000
MXE	MROH7	chr1	+	54682655	54682794	0.020226	1.000000
3'ASS	KHDC4	chr1	-	155916624	155916737	0.020291	1.000000
3'ASS	ZC3H12A	chr1	+	37482349	37482540	0.020432	1.000000
RI	SLC2A6	chr9	-	133478253	133479096	0.020473	0.520593
5' ASS	KIF12	chr9	-	114097518	114097741	0.020484	1.000000
RI	RELA	chr11	-	65661936	65662205	0.020534	0.520593
RI	TNFAIP2	chr14	+	103129739	103130124	0.020682	0.520593
RI	ARHGEF40	chr14	+	21080872	21082119	0.020722	0.520593
SE	PHF20L1	chr8	+	132825034	132825157	0.020857	1.000000
RI	SELENBP1	chr1	-	151365767	151366453	0.020943	0.522759
5' ASS	TMEM134	chr11	-	67468016	67468092	0.021150	1.000000
3'ASS	UBL4A	chrX	-	154485770	154486083	0.021245	1.000000
RI	PITPNM1	chr11	-	67496177	67497436	0.021254	0.527165
SE	TOP3B	chr22	-	21975980	21976187	0.021306	1.000000
RI	NPAS2	chr2	+	100982230	100988276	0.021456	0.528793
RI	TTC31	chr2	+	74490243	74491184	0.021696	0.530537
RI	HDAC6	chrX	+	48814674	48814893	0.021802	0.530537
3'ASS	RAP1GAP	chr1	-	21597960	21598095	0.021807	1.000000
3'ASS	CENPV	chr17	-	16348615	16349333	0.021831	1.000000
RI	CROCCP3	chr1	-	16476929	16478437	0.021994	0.530537
RI	MTA1	chr14	+	105450057	105450324	0.022072	0.530537
SE	FOXP2	chr7	+	114570806	114570881	0.022214	1.000000
RI	ABCA2	chr9	-	137022701	137023052	0.022333	0.533528
SE	TARBP2	chr12	+	53502014	53502184	0.022526	1.000000
3'ASS	FAM122B	chrX	-	134769566	134772283	0.022549	1.000000
RI	RAD54L2	chr3	+	51630271	51630931	0.022577	0.534902
3'ASS	ZMYND8	chr20	-	47298728	47298947	0.022629	1.000000

5' ASS	MCMBP	chr10	-	119843247	119843426	0.022644	1.000000
RI	PDXDC1	chr16	+	15031734	15032979	0.022665	0.534902
SE	HAUS4	chr14	-	22951554	22951689	0.023068	1.000000
5' ASS	CLMN	chr14	-	95194214	95194596	0.023254	1.000000
3'ASS	SLC25A29	chr14	-	100295625	100295993	0.023272	1.000000
SE	ZNF655	chr7	+	99571687	99571792	0.023278	1.000000
RI	YIF1A	chr11	-	66284579	66284966	0.023376	0.545320
SE	GRB10	chr7	-	50732271	50732368	0.023497	1.000000
RI	C21orf58	chr21	-	46314870	46315547	0.023501	0.545320
RI	ADCY6	chr12	-	48770765	48771973	0.023527	0.545320
5' ASS	CFAP410	chr21	-	44334060	44335804	0.023635	1.000000
3'ASS	VEZF1	chr17	-	57980602	57980804	0.023734	1.000000
SE	SSBP4	chr19	+	18431352	18431418	0.023790	1.000000
5' ASS	MRPL55	chr1	-	228108436	228109021	0.023808	1.000000
RI	WDR90	chr16	+	653745	655147	0.023829	0.549048
RI	TNS1	chr2	-	217880897	217882411	0.024133	0.552789
RI	BAZ2A	chr12	-	56604584	56605327	0.024381	0.553077
RI	MED25	chr19	+	49836225	49836979	0.024478	0.553077
SE	SYNE2	chr14	+	64215285	64215354	0.024515	1.000000
RI	P2RX2	chr12	+	132621474	132621972	0.024641	0.553077
RI	TMEM25	chr11	+	118531191	118531871	0.024714	0.553077
RI	B4GALNT3	chr12	+	552466	553983	0.024952	0.555226
5' ASS	THOC5	chr22	-	29543414	29543542	0.025010	1.000000
SE	ACIN1	chr14	-	23067313	23068671	0.025177	1.000000
3'ASS	GATD1	chr11	-	770992	771212	0.025443	1.000000
SE	RICTOR	chr5	-	38949355	38949427	0.025843	1.000000
RI	TAF1C	chr16	-	84183075	84183328	0.025878	0.572545
RI	TMEM184A	chr7	-	1548518	1549945	0.026070	0.573536
SE	NMI	chr2	-	151288890	151289077	0.026295	1.000000
5' ASS	TSC1	chr9	-	132925586	132927304	0.026315	1.000000
RI	PPOX	chr1	+	161167099	161167486	0.026565	0.581145
SE	DGUOK	chr2	+	73938909	73939022	0.026584	1.000000
SE	CPEB4	chr5	+	173932449	173932500	0.026647	1.000000
5' ASS	PACRGL	chr4	+	20700457	20700747	0.026689	1.000000
3'ASS	NDUFA3	chr19	+	54106763	54106961	0.026749	1.000000
RI	THOC6	chr16	+	3026062	3026288	0.026754	0.582016
RI	GRIA2	chr4	+	157334009	157335877	0.026992	0.583923
5' ASS	CUL2	chr10	-	35090178	35090642	0.027032	1.000000
3'ASS	LETMD1	chr12	+	51053777	51053860	0.027171	1.000000
RI	SH3TC1	chr4	+	8235432	8237670	0.027427	0.586660
RI	MAPKBP1	chr15	+	41812515	41813781	0.027516	0.586660

SE	DIS3L2	chr2	+	232198526	232198776	0.027538	1.000000
5' ASS	ARHGAP26	chr5	+	143014079	143014293	0.027713	1.000000
RI	AKNA	chr9	-	114343707	114346009	0.027743	0.586660
RI	PER1	chr17	-	8146896	8147381	0.027760	0.586660
RI	DPP7	chr9	-	137114242	137114576	0.027872	0.586660
SE	HNRNPAB	chr5	+	178210131	178210272	0.027909	1.000000
SE	EMC10	chr19	+	50481860	50481947	0.027914	1.000000
3'ASS	FBXW9	chr19	-	12689961	12690200	0.027923	1.000000
3'ASS	SUPT20H	chr13	-	37022010	37023044	0.028066	1.000000
SE	RGS14	chr5	+	177371474	177371589	0.028226	1.000000
SE	BTD	chr3	+	15632590	15632866	0.028489	1.000000
3'ASS	LSR	chr19	+	35266835	35266967	0.028729	1.000000
SE	REPIN1	chr7	+	150370759	150370884	0.028827	1.000000
RI	ABCF3	chr3	+	184187396	184187761	0.028935	0.605772
SE	TLDC1	chr16	-	84502720	84502862	0.029025	1.000000
3'ASS	STARD3	chr17	+	39656909	39657085	0.029261	1.000000
SE	TCF12	chr15	+	57252420	57252492	0.029352	1.000000
SE	DPP9	chr19	-	4715528	4715726	0.029399	1.000000
SE	KCTD21	chr11	-	78177738	78177985	0.029497	1.000000
SE	MFF	chr2	+	227328677	227330846	0.029535	1.000000
RI	VPS28	chr8	-	144424214	144424819	0.029679	0.615342
RI	COQ8B	chr19	-	40705095	40705447	0.029708	0.615342
SE	LLGL2	chr17	+	75538530	75538636	0.029738	1.000000
5' ASS	ODF2	chr9	+	128460941	128461067	0.029794	1.000000
SE	ZNF711	chrX	+	85246929	85247183	0.030199	1.000000
5' ASS	MYO15B	chr17	+	75624542	75624688	0.030275	1.000000
SE	ARHGAP12	chr10	-	31839636	31839711	0.030292	1.000000
RI	PDE1B	chr12	+	54570240	54572741	0.031380	0.643698
RI	SPATA20	chr17	+	50549984	50550312	0.031408	0.643698
MXE	MFSD8	chr4	-	127939852	127939997	0.031503	1.000000
SE	APCDD1L-DT	chr20	+	58519503	58519599	0.031725	1.000000
MXE	GIT2	chr12	-	109947255	109947504	0.031748	1.000000
RI	MICALL2	chr7	-	1436741	1437608	0.031985	0.652101
SE	AFTPH	chr2	+	64579485	64581273	0.032386	1.000000
SE	BANF1	chr11	+	66002843	66003070	0.032652	1.000000
5' ASS	RNF123	chr3	+	49712478	49712892	0.032702	1.000000
SE	RAB40B	chr17	-	82696388	82696853	0.032767	1.000000
RI	CPNE2	chr16	+	57146084	57147964	0.032820	0.658808
5' ASS	ZFPL1	chr11	+	65085114	65085314	0.032907	1.000000
RI	RNF123	chr3	+	49704649	49705186	0.032963	0.658808
RI	PTPN21	chr14	-	88469498	88470050	0.032969	0.658808

RI	TOP3B	chr22	-	21957026	21958693	0.032991	0.658808
RI	VPS13D	chr1	+	12348822	12349374	0.033402	0.661686
RI	CLN3	chr16	-	28491481	28491835	0.033475	0.661686
SE	TRMO	chr9	-	97916163	97916338	0.033841	1.000000
SE	FOXK2	chr17	+	82595781	82596191	0.034324	1.000000
SE	GTF3C5	chr9	+	133054943	133054964	0.034393	1.000000
3'ASS	GGA2	chr16	-	23493359	23493587	0.034436	1.000000
RI	LPCAT4	chr15	-	34359588	34360209	0.034477	0.678048
SE	ATP2C1	chr3	+	130894730	130894775	0.034502	1.000000
RI	MPDZ	chr9	-	13133823	13136182	0.034893	0.682783
RI	FBXL6	chr8	-	144356989	144357502	0.035277	0.684394
5' ASS	SMG9	chr19	-	43739672	43740218	0.035317	1.000000
RI	NSUN5P1	chr7	+	75412788	75415023	0.035327	0.684394
SE	MTO1	chr6	+	73480292	73480367	0.035330	1.000000
SE	SLC2A11	chr22	+	23877384	23877869	0.035698	1.000000
RI	MPDU1	chr17	+	7586898	7587271	0.035768	0.689251
3'ASS	SH3TC1	chr4	+	8236277	8237670	0.035854	1.000000
5' ASS	TIA1	chr2	-	70216752	70216994	0.035909	1.000000
RI	CDK5RAP3	chr17	+	47974399	47975337	0.035932	0.689251
5' ASS	PTPN4	chr2	+	119915178	119916690	0.036109	1.000000
SE	ZNF415	chr19	-	53115209	53115307	0.036256	1.000000
3'ASS	ANKRD18A	chr9	-	38571357	38573129	0.036516	1.000000
SE	SEC23B	chr20	+	18508090	18508297	0.036567	1.000000
RI	SLC5A6	chr2	-	27202812	27203345	0.036816	0.694557
RI	L3MBTL2	chr22	+	41227088	41227869	0.036825	0.694557
RI	GTF2H4	chr6	+	30913109	30913387	0.036936	0.694557
RI	ACADVL	chr17	+	7222666	7223237	0.036991	0.694557
RI	DGKA	chr12	+	55939414	55940433	0.037100	0.694557
RI	NELFE	chr6	-	31954554	31955096	0.037677	0.701985
SE	RAB11FIP3	chr16	+	491133	491268	0.037935	1.000000
3'ASS	LRRC29	chr16	-	67210047	67211024	0.037939	1.000000
3'ASS	USP19	chr3	-	49115723	49115944	0.037970	1.000000
3'ASS	LRRC29	chr16	-	67210047	67211073	0.037972	1.000000
5' ASS	UMPS	chr3	+	124737567	124738443	0.038104	1.000000
RI	TMEM94	chr17	+	75492928	75493593	0.038281	0.709833
SE	NOP58	chr2	+	202295673	202295837	0.038498	1.000000
RI	RPS6KB2	chr11	+	67433125	67433447	0.038814	0.714509
RI	PAF1	chr19	-	39390068	39390289	0.038963	0.714509
RI	SFI1	chr22	+	31604868	31606430	0.039271	0.714509
RI	LRCH4	chr7	-	100577085	100577389	0.039422	0.714509
RI	ABHD17A	chr19	-	1877507	1880115	0.039450	0.714509

RI	GPS1	chr17	+	82056839	82057470	0.039695	0.715604
3'ASS	FOXP4	chr6	+	41589962	41590170	0.039990	1.000000
SE	CEP164	chr11	+	117335604	117335680	0.039996	1.000000
SE	C12orf65	chr12	+	123233739	123233924	0.040469	1.000000
RI	DDX11	chr12	+	31103576	31104791	0.040514	0.727004
SE	RPH3AL	chr17	-	333758	333934	0.040571	1.000000
RI	MRPL52	chr14	+	22829886	22830110	0.040963	0.728616
RI	LY6H	chr8	-	143158802	143159709	0.040978	0.728616
SE	TMEM263	chr12	+	106957081	106957149	0.041237	1.000000
RI	PISD	chr22	-	31620995	31621472	0.041350	0.731889
RI	METTL17	chr14	+	20995900	20996292	0.041563	0.732332
3'ASS	ENOSF1	chr18	-	690548	690745	0.041870	1.000000
RI	POFUT2	chr21	-	45263927	45267713	0.041992	0.734270
RI	ZNF34	chr8	-	144778037	144778525	0.042050	0.734270
SE	DDX47	chr12	+	12821207	12821396	0.042286	1.000000
RI	PFAS	chr17	+	8256266	8256648	0.042436	0.737708
SE	ATG16L2	chr11	+	72823543	72823797	0.042461	1.000000
SE	TMCC1	chr3	-	129860939	129861012	0.042632	1.000000
3'ASS	SIRT2	chr19	-	38893818	38893971	0.042687	1.000000
MXE	MFSD8	chr4	-	127938782	127938838	0.042732	1.000000
5' ASS	ANXA4	chr2	+	69810593	69811456	0.042794	1.000000
RI	TRPT1	chr11	-	64224284	64224717	0.042839	0.741405
SE	MRPL55	chr1	-	228108436	228109021	0.043216	1.000000
3'ASS	TMEM268	chr9	+	114624135	114624459	0.043365	1.000000
RI	RTEL1-TNFRSF6B	chr20	+	63688527	63689132	0.043825	0.755109
SE	MCTP2	chr15	+	94278153	94278263	0.044021	1.000000
5' ASS	MADD	chr11	+	47295495	47295579	0.044054	1.000000
RI	GPATCH4	chr1	-	156596080	156596478	0.044548	0.764189
SE	DDR1	chr6	+	30885624	30885680	0.044973	1.000000
RI	LUC7L	chr16	-	227241	228402	0.045340	0.774367
5' ASS	IKBKB	chr8	+	42293442	42293568	0.045534	1.000000
SE	TBC1D17	chr19	+	49879322	49879504	0.045648	1.000000
RI	NAT9	chr17	-	74771959	74772277	0.045785	0.775691
5' ASS	NFKBIL1	chr6	+	31557627	31557849	0.045794	1.000000
RI	MPP2	chr17	-	43882902	43883355	0.045816	0.775691
SE	MRPL55	chr1	-	228108436	228109021	0.046045	1.000000
SE	PRKDC	chr8	-	47782161	47782254	0.046673	1.000000
SE	ZNRD1	chr6	+	30061916	30062017	0.046681	1.000000
RI	RPS16	chr19	-	39433521	39433761	0.046748	0.788031
RI	BBS1	chr11	+	66515539	66515731	0.047312	0.794109

SE	MPHOSPH9	chr12	-	123227462	123227616	0.047850	1.000000
MXE	MEG3	chr14	+	100832511	100832641	0.047995	1.000000
5' ASS	ITFG2	chr12	+	2818105	2818353	0.048003	1.000000
MXE	FRMD8	chr11	+	65389360	65389528	0.048164	1.000000
SE	MAPK9	chr5	-	180279795	180280036	0.048175	1.000000
3'ASS	SLC25A10	chr17	+	81716680	81716855	0.048270	1.000000
RI	ZWINT	chr10	-	56360017	56360383	0.048350	0.804746
RI	PHGDH	chr1	+	119741766	119743044	0.048359	0.804746
3'ASS	TMEM175	chr4	+	951645	951717	0.048453	1.000000
SE	SDHAP1	chr3	-	195984593	195984737	0.048883	1.000000
RI	PKD1	chr16	-	2105321	2106024	0.049098	0.811410
SE	ALDOA	chr16	+	30065875	30065927	0.049107	1.000000
RI	PPIL2	chr22	+	21694936	21695496	0.049176	0.811410
SE	ABHD11	chr7	-	73737220	73737391	0.049476	1.000000
SE	ZNF655	chr7	+	99561887	99561973	0.049560	1.000000
SE	YPEL5	chr2	+	30148244	30148541	0.049587	1.000000
SE	SYNRG	chr17	-	37577379	37577613	0.049855	1.000000
3'ASS	SCRN2	chr17	-	47839443	47839858	0.049966	1.000000